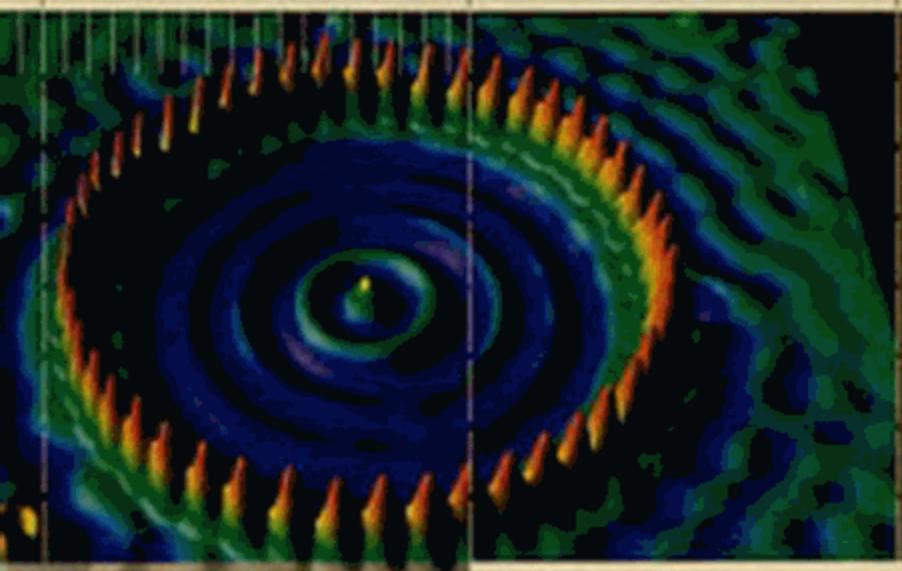


物理系列

The Quantum Quark

量子夸克

[英] 安德鲁·华生 / 著　刘 健　雷奕安 / 译

第一推动

湖南科学技术出版社

The Quantum Quark by Andrew Watson
© Andrew Watson 2004
All rights reserved including the rights of reproduction in whole or in part in any form. Simplified Chinese edition copyright: 2006 Hunan Science & Technology Press
湖南科学技术出版社通过剑桥大学出版社独家获得本书中文简体字版中国大陆地区出版发行权
本书根据剑桥大学出版社2004年版本译出
著作权合同登记号：18-2006-036
版权所有，侵权必究

图书在版编目（CIP）数据
量子夸克 / （英）华生著；刘健，雷奕安译.—长沙：湖南科学技术出版社，2008.4
(第一推动丛书)
书名原文:The Quantum Quark
ISBN 978-7-5357-5236-9
I. 量… II.①华…②刘…③雷… III.夸克—普及读物 IV.O572.3-49
中国版本图书馆 CIP 数据核字（2008）第 049993 号

第一推动丛书　物理系列
量子夸克
著　　者：[英]安德鲁·华生
译　　者：刘　健　雷奕安
责任编辑：吴　炜　戴　涛
出版发行：湖南科学技术出版社
社　　址：长沙市湘雅路 276 号
http://www.hnstp.com
邮购联系：本社直销科　0731 - 4375808
印　　刷：长沙化勘印刷有限公司
（印装质量问题请直接与本厂联系）
厂　　址：长沙市青园路 4 号
邮　　编：410004
出版日期：2008 年 4 月第 1 版第 1 次
开　　本：880mm×1230mm　1/32
印　　张：16.25
字　　数：396000
书　　号：ISBN 978-7-5357-5236-9
定　　价：39.50 元
（版权所有　翻印必究）

总　序

科学，特别是自然科学，最重要的目标之一，就是追寻科学本身的原动力，或曰追寻其第一推动。同时，科学的这种追求精神本身，又成为社会发展和人类进步的一种最基本的推动。

科学总是寻求发现和了解客观世界的新现象，研究和掌握新规律，总是在不懈地追求真理。科学是认真的、严谨的、实事求是的，同时，科学又是创造的。科学的最基本态度之一就是疑问，科学的最基本精神之一就是批判。

的确，科学活动，特别是自然科学活动，比较起其他的人类活动来，其最基本特征就是不断进步。哪怕在其他方面倒退的时候，科学却总是进步着，即使是缓慢而艰难地进步，这表明，自然科学活动中包含着人类的最进步因素。

正是在这个意义上，科学堪称为人类进步的“第一推动”。

科学教育，特别是自然科学的教育，是提高人们素质的重要因素，是现代教育的一个核心。科学教育不仅使人获得生活和工作所需的知识和技能，更重要的是使人获得科学思想、科学精神、科学态度以及科学方法的熏陶和培养，使人获得非生物本能

的智慧，获得非与生俱来的灵魂。可以这样说，没有科学的“教育”，只是培养信仰，而不是教育。没有受过科学教育的人，只能称为受过训练，而非受过教育。

正是在这个意义上，科学堪称为使人进化为现代人的“第一推动”。

近百年来，无数仁人志士意识到，强国富民再造中国离不开科学技术，他们为摆脱愚昧与无知作了艰苦卓绝的奋斗。中国的科学先贤们代代相传，不遗余力地为中国的进步献身于科学启蒙运动，以图完成国人的强国梦。然而应该说，这个目标远未达到。今日的中国需要新的科学启蒙，需要现代科学教育。只有全社会的人具备较高的科学素质，以科学的精神和思想、科学的态度和方法作为探讨和解决各类问题的共同基础和出发点，社会才能更好地向前发展和进步。因此，中国的进步离不开科学，是毋庸置疑的。

正是在这个意义上，似乎可以说，科学已被公认是中国进步所必不可少的推动。

然而，这并不意味着，科学的精神也同样地被公认和接受。虽然，科学已渗透到社会的各个领域和层面，科学的价值和地位也更高了，但是毋庸讳言，在一定的范围内，或某些特定时候，人们只是承认“科学是有用的”，只停留在对科学所带来的后果的接受和承认，而不是对科学的原动力、科学的精神的接受和承认。此种现象的存在也是不能忽视的。

科学的精神之一，是它自身就是自身的“第一推动”。也就是说，科学活动在原则上是不隶属于服务于神学的，不隶属于服务于儒学的，科学活动在原则上也不隶属于服务于任何哲学。科学是超越宗教差别的，超越民族差别的，超越党派差别的，超越文化的地域的差别的，科学是普适的、独立的，它自身就是自身的主宰。

湖南科学技术出版社精选了一批关于科学思想和科学精神的世界名著，请有关学者译成中文出版，其目的就是为了传播科学的精神，科学的思想，特别是自然科学的精神和思想，从而起到倡导科学精神，推动科技发展，对全民进行新的科学启蒙和科学教育的作用，为中国的进步作一点推动。丛书定名为《第一推动》，当然并非说其中每一册都是第一推动，但是可以肯定，蕴含在每一册中的科学的内容、观点、思想和精神，都会使你或多或少地更接近第一推动，或多或少地发现，自身如何成为自身的主宰。

《第一推动丛书》编委会

前　言

掂一下这本书，你就能感觉到地心引力的作用。看到这一页字，你会意识到光的存在：引力和电磁力对我们来说，都是再熟悉不过的东西了。但是大自然还操纵着另外两种我们不怎么熟悉的力。要了解这两种力的机理，就要在进行高倍放大后，在特别精细的尺度上观察世界，这时日常事物都会呈现出更为基本的组成部分。这两种我们不太熟悉的力，分别叫做强相互作用力（strong force，简称强力）和弱相互作用力（weak force，简称弱力）。强力使组成物质的最基本成分结合在一起，为我们的太阳提供动力，而弱力则参与放射性过程。物理学家们已经提出了一套理论，来说明强力怎样在最基本的层次上产生作用，本书的内容就是关于这套理论的。这套理论就是量子色动力学（quantum chromodynamics），或简称为 QCD。它是人类智慧最伟大的成就之一。

研究强相互作用力的物理学分支叫做粒子物理，或者也可以叫做高能物理。高能物理学家也研究弱相互作用力，但是我们可以耍个小把戏，把有关弱力的东西，连同包括引力在内的所谓大

统一理论都放到一边。就像一本历史书，可能仅涵盖了英国历史，而不是整个欧洲的历史一样，本书的主要目的是介绍量子色动力学及其基本规律和发展历程。就像一本讲述英国历史的书并不能记载所有的英国历史一样，本书也不可能面面俱到。为了方便更广大的读者，与其他介绍粒子物理的书相比，本书更加详细地探讨了量子色动力学及其相关内容。

《量子夸克》一书意在使广大读者容易理解书中内容：读者不需要有物理学或数学基础，本书使用的符号也做了最大程度的简化。如同一本俄国名著中会有许多人物的名字令我们感到陌生那样，本书也会有很多晦涩的术语，因此书后提供了详尽的术语表供大家查阅。书后还收录了重大事件的历史年表，以帮助读者勾勒出 QCD 随时间发展的脉络。量子色动力学是一个庞大的课题，许多人为此投入了大量时间和精力，如此多的内容无法全部包罗在本书当中。我希望有关专家们能原谅我略去了一些观点、人物、实验以及参考资料，权且让我们在伟大思想之间信步，做一次轻松的旅行，而不要把本书作为一本教科书来对待。

致 谢

本书的完成离不开许多人的大力支持，太多的人需要感谢，我在这里无法一一列举。但我还是要在这里对一些人示以谢意。我特别感谢 Tracey 和 Mark Williamson，还有 Graham Keeley，他们在我写这本书的早期为我提供了各种帮助。Arturo Sangalli 和 Rob Penfold 自始至终给了我很多支持和鼓励。剑桥大学的布莱恩·韦伯（Bryan Webber）和卡文迪许实验室（Cavendish Laboratory）为我提供了所需资料。我还要感谢几位匿名评论专家的帮助，他们找出了我早期书稿中的一些不足之处。还要特别感谢 Elke-Caroline Aschenauer，Christine Davies，Gordon Kane，Donald Perkins，Alexander Polyakov 和 Mikhail Shifman，他们花了大量的时间检查了全书或部分内容，并提供了许多宝贵的意见和建议。由于本人水平所限，书中不免还有一些纰漏以及不准确之处，还望大家谅解和指正。最后，我衷心感谢那些将自己多年珍藏的图片无私提供给我的科学家们。

目 录

前言 …… 1

致谢 …… 1

第一章 力及其相关理论简介 …… 1

第二章 对称性 …… 12

第三章 量子世界 …… 24

第四章 向 QCD 进军 …… 129

第五章 QCD 理论的一个数字 …… 318

第六章 群居的胶子 …… 341

第七章 夸克和强子 …… 361

第八章 显微镜下的夸克 …… 387

第九章 庸人自扰 …… 434

第十章 棋盘 QCD 理论 …… 458

附录 1 QCD 年表 …… 482

附录 2 小辞典 …… 492

第一章 力及其相关理论简介

在路易斯安那州（Louisiana）的利文斯通（Livingston），坐落着一幢与众不同的“L”形建筑物。它由两条相互垂直的 4 千米长的臂组成，两臂的远端和两臂连接的拐角处都装有反射镜，激光经镜子反射后在两臂中来回穿梭。在往返多次之后，光线会重新汇合①。测量结果令人吃惊：建筑物的长度发生了微小的改变。

建筑物的长度为什么会改变呢？难道路易斯安那州自身发生了伸缩？科学家们普遍赞同这个想法。也就是说，如果艾尔伯特·爱因斯坦（Albert Einstein）的引力理论正确，引力波会穿过路易斯安那州，同样也会经过华盛顿州（Washington state）。在华盛顿州的汉福德②，也有与路易斯安那州一模一样的仪器正在观测这种波动。波动会穿过整个地球。实际上，根据爱因斯坦的广义相对论，波动将穿过时空结构本身。只不过在路易斯安那州和华盛顿州能方便地观察到这种波动而已。

引起波动的振动源非常强，能够使整个宇宙都发生震颤，诸

---

① 译者注：这是利用激光干涉测量长度。

② 译者注：Hanford，美国华盛顿州南部重要的原子能研究中心。

如恒星的坍塌，甚至是宇宙诞生本身都属于这种振动源。广义相对论预言，这些事件将会发出引力波，这就是科学家们在路易斯安那州和华盛顿州想要探测的那种时空波动。在意大利、澳大利亚等其他一些地方也有探测器在探测这种波动。

广义相对论还预言了黑洞（black hole，一种坍塌后的恒星）的存在，其强大的引力连光都无法挣脱。天文学家们无法直接想象黑洞的样子，但他们已经发现了许多有可能是黑洞的天体，像天鹅座 X1 恒星系统（Cygnus X-1 system）中的伴星就极有可能是黑洞。天文学家认为，我们银河系的中心也隐藏着一个巨大的黑洞。他们的观测也支持了广义相对论的另一个预言——引力透镜效应（gravitational lensing）。就是说，如果在地球上观测一个很大的星系，可能会看到星系的多个像，这些像会按照一定的方式排列，并且看起来可能比实际位置更远。爱因斯坦最初的想法是，太阳会使恒星发出的光弯曲，引力透镜效应是这种想法的另一个说法，人们第一次观测到引力透镜效应是在 1979 年①。而太阳会使光线弯曲的预言，早在 1919 年就已经被证实②，正是这一成就使爱因斯坦和广义相对论登上了历史舞台。

爱因斯坦理论的出发点，是人们在电梯中得到的日常经验。当电梯开始上升的时候，你能感觉到自己变重了；当电梯开始下降，你会觉得自己变轻了。根据爱因斯坦的想法，引力产生的效果同加速产生的效果是无法区分的。他把这个原理同一种关于弯曲空间的新兴数学联系到一起，得出了一套关于引力的理论。在这套理论中，万有引力源自于质量引起的的时空扭曲。他的广义相对论取代了艾萨克·牛顿（Isaac Newton）在 17 世纪提出的引

---

① 译者注：1979 年，Walsh，Carswell 和 Weymann 发现了首个类星体 0957+561 的引力透镜现象。

② 译者注：1919 年，爱丁顿率领的英国科考队对日食的观测验证了广义相对论的这一推论。

力定律。牛顿的理论只是引力理论的一种特例，只适用于处理诸如苹果落地和行星运动等现象。

万有引力虽然在处理宇宙尺度上的问题时极其出色，但在极小的尺度上却变得有名无实。尽管具有质量的物质之间普遍存在引力，这对于物质世界成员中最微小的基本粒子也不例外，但是引力作用在距离很小的情况下被自然界的其他 3 种力完全淹没了。并且到目前为止，没有人能拼凑出一套合适的理论，用来解释引力如何在极小尺度范围内作用。所以，我们在讲述量子色动力学的过程中把引力忽略掉并没有多大的影响。

无处不在的电

曾经有个小故事，讲的是时任英国财政大臣的威廉姆·格莱斯顿（William Gladstone），问发电机和电动机的发明者米歇尔·法拉第（Michael Faraday）："电将会有什么实际用途?"法拉第已经认识到自己的发现将会引起重大的社会变革，他回答说："先生，总有一天您会向人民征收电税的。"

从那时起，人类已经在征服电和磁（两者共同构成了电磁作用）的道路上取得了长足的进步。譬如现在，几乎在全球的每个角落都能够看到流行的肥皂剧，而这一切都离不开电磁学的发展：正是法拉第式的发电机提供了大部分电能，信号通过被称为无线电波的电磁波传播，电磁场控制着电视机中形成图像的电子束。

继引力之后，电磁力成为第二种在科学家面前揭开神秘面纱的力。电荷是电现象的源头，而磁效应的产生，也总可以追溯到电荷的运动。在法拉第发现的指引下，詹姆士·克拉克·麦克斯韦（James Clerk Maxwell）在他的电磁理论中成功地统一了电和磁。他的理论反过来又引发了这样的想法：电磁波就是振动着的

电波和磁波，它们以固定的速度一起传播。整个电磁波谱的范围从低频端的无线电波过渡到可见光，一直延伸到高频端的 X 射线。也就是说，光也是一种电磁波。

电动机、发电机以及无线电波等，都属于被物理学家们称作“经典电磁学”的学科所研究的范畴。之所以称作“经典”，是为了把它与 20 世纪初期诞生的新电磁学相区别。那时人们已经意识到，与之前的想法不同，光以及其他一些物质都不是简单地均匀分布的。实际上按照这个观点，光有时看上去像是由粒子组成的。在极其细微的尺度——原子尺度上，两个带电物体之间的相互作用，是通过交换一种分立的电磁载体实现的。这种电磁载体叫做光子（photon），用符号 γ 表示。从本质上来说，微小的带电粒子就是通过交换“光粒子”（即光子）相互作用的。

光子每时每刻都围绕着我们。为简化起见，我们用指数形式表示数字，100 可写成 10^2，1000 可以写作 10^3，那么，一个家用白炽灯泡每秒钟能发出大约 10^{20} 个光子。光子数目如此庞大，而每个光子的能量又是如此的小，根据日常的生活经验，人们总是把这些光子作为一个整体来处理。但是在单个原子尺度上，光与物质粒子之间以及带电的物质粒子之间的电磁相互作用，都是由光的粒子本性决定的。

构建现有的微观电磁学理论是一项浩大的工程，这项工程从 1927 年开始启动，直到 1950 年左右才全部竣工。该理论被称作量子电动力学或 QED（Quantum Electrodynamics），它描述了微观世界中光子与物质粒子作用的一般规律。这方面的成果并不是本书将要叙述的内容，本书要介绍的是有关强相互作用力（以下简称强力）的理论，叫做量子色动力学或 QCD（Quantum Chromodynamics）。但是 QED 比 QCD 更简单，并且 QED 的提出比 QCD 早了几十年，它为 QCD 的发展提供了一套基本的模式。从这种意义上来说，了解光与物质粒子之间的相互作用，是理解

QCD 的一种途径。

亚原子探索

芝加哥、日内瓦、斯坦福、汉堡、东京——以上所列举的是世界上最大的几座粒子加速器的所在地。这些分布在全球各地的粒子加速器（除了上文提到的几座外，还有许多加速器分布在其他地方），形成了物理实验的前沿阵地，科学家们期望能通过它们揭示自然界中另外两种力的秘密。这项任务异常艰巨，相关工作量大得惊人，耗费的资源也相当多。除了主要的实验室之外，全世界还有共 6000 多个研究机构和大学在分析实验结果、构建相关理论，为这些实验提供理论支持。总计有上万人全身心地投入到这项探索事业中来，投入的资金达上百亿美元。

那么他们到底在寻找什么呢？化学家们公认，组成物质的可辨认的最小微粒是原子（atom），原子内部几乎全是空的。简单地说，原子大概是个模糊的球状物，内部包含一种带负电的粒子，叫做电子（electron），还有一个带正电的坚硬核心部分，叫原子核（nucleus）。一个原子的直径大约是 10^{-10} 米（这里我们用 10^{-2} 来标记 1/100，用 10^{-3} 表示 1/1000）。也就是说，100 亿个原子的直径加起来才 1 米长。原子核的直径仅是原子直径的十万分之一，也就是大约 10^{-15} 米，但它却占据了原子约 99.975%的质量。

欧内斯特·卢瑟福（Ernest Rutherford）在 1911 年证明了原子核的存在，并在几年之后，猜出原子核内应该有一种叫做质子（proton）的粒子。詹姆士·查德威克（James Chadwick）在 1932 年发现了原子核的另一种组成部分——中子（neutron）。一个新的物质层次逐渐呈现出来。

质子的质量很小，大约为 1.673×10^{-27} 千克，带正电荷，且

电量与电子所带的负电荷相等。质子通常用符号 p 表示。最简单的元素——氢元素的原子由 1 个电子和只含有 1 个质子的原子核组成。中子的质量与质子质量差不多，大约为 1.675×10^{-27} 千克。中子质量与质子质量如此接近，但又不完全相同，这并不是巧合。中子通常用字母 n 表示，不带电荷。除氢之外最简单的元素当属氦了，宴会上那些能飘在空中的气球，一般都是用氦气吹起来的。氦原子的原子核包含 2 个质子和 2 个中子：原子核中的质子数与中子数都大致相等，而原子中的电子数目，总是等于其原子核中的质子数。

质子和中子是 QCD 理论中的主角，它们会在以后经常出现。我们已经顺便介绍过电子了，电子通常用符号 e 表示，质量只有 9.11×10^{-30} 千克，或者说大约是质子或中子质量的1/2000。它所带的电荷是负的，并且电量非常小，约为 -1.602×10^{-19} 库仑。每秒钟内流经家用电缆的电荷所带的电量，大约是电子电量的 10^{19} 倍。到目前为止，大家已经承认电子是真正的基本粒子，它不可分割，并且目前仍没测量出电子的大小。

但质子和中子却与电子不同，这正是所有问题的关键所在。还有另外两种力决定着中子和质子内部的结构，以及质子和中子是怎样结合在一起的。核物理学家们对质子和中子怎样“粘”在一起形成原子核更感兴趣。粒子物理学家们则会研究更深层次的问题，他们主要探讨质子和中子本身的内部结构以及亚原子核世界中力的基本性质。研究人员就是在用那些巨大的粒子加速器做这些事情。

揭开弱力的面纱

时间回溯到 1983 年，在瑞士日内瓦，欧洲粒子物理实验室（CERN，最初的完整名称为 Centre Européen pour la Recherche

Nucléaire）的科学家们做出了20世纪一项重大的科学发现。

日内瓦地下有一座周长将近7千米的圆形隧道，隧道中放置着环绕强大磁场的真空管道。磁场是用来引导带电粒子沿管道运动的，它能使带电粒子按管道的形状运动而不会碰到管壁，并且能使粒子在运动时聚集成一群。每当这些粒子经过时，巨大的射频发生器就会有节奏地把脉动的能量传输给粒子，加快粒子的速度，粒子最后能以不可思议的速度在管道中运动。在同一次实验中，管道中存在两种类型的粒子：质子和质子的反粒子——反质子，反质子用符号 $\bar{p}$ 表示。反质子的性质稍候再做详细介绍，现在只要知道，它与质子有相同的质量，但带有相反的电荷就足够了。在磁场的作用和射频脉冲的加速下，反质子在管道中与质子朝着相反的方向，一圈一圈地做着圆周运动。

实验需要格外小心，要确保在绝大多数时间内，这两束反方向做圆周运动的粒子束不会碰面。然而在某些情况下，又需要粒子束在磁场的控制下被迫相交。质子和反质子发生碰撞，双方都撞得粉身碎骨，并向安放在粒子束交汇区域四周的大型探测器发出辐射。这些探测器中的每一台都有一幢房子那么大，但比房子要重得多，内部有多个辐射探测层，用来监视是否有从碰撞点发出的粒子经过。

这种辐射同时发出多种粒子。运算能力强大的计算机阵列会分析探测信号，重新绘制出通过探测器的粒子轨迹，并且计算出粒子的质量和所带电荷。就是在这里，人们在实验中第一次发现了W粒子，此后不久，又第一次观测到了Z粒子的存在。

两种W粒子（指带正电荷的 W^+ 粒子和带负电荷的 W^- 粒子）和不带电的Z粒子的发现者为意大利物理学家卡罗·鲁比亚（Carlo Rubbia）和荷兰人西蒙·范德·米尔（Simon van der Meer），他们赢得了1984年的诺贝尔物理学奖。

实际上，理论物理学家谢尔登·格拉肖（Sheldon Glashow）、

阿卜杜勒·萨拉姆（Abdus Salam）和史蒂芬·温伯格（Steven Weinberg）在此之前就提出了一套基于这些 W 粒子和 Z 粒子的理论，这套理论也使他们获得了诺贝尔奖①。他们的理论有一个结论，可以严格检验理论本身是否正确，而这一结论可以在实验中观察到：W 粒子的质量与 Z 粒子的质量必须满足一定的关系。此外，这两种粒子的质量还都与一个于 1973 年首次在实验中测得的量有关。所以理论物理学家们在鲁比亚建造他的探测器之前，就知道 W 粒子和 Z 粒子的质量应当是多少。这就是为什么发现 W 粒子和 Z 粒子，并且观测到它们具有理论预期的质量，是一件如此意义重大和激动人心的事情。现在，加速器可以大批量生产这些粒子，为实验物理学家们仔细地测量它们的性质提供了便利条件。除了一个令人感到困惑的问题还悬而未决之外，一切都十分令人满意。

格拉肖-温伯格-萨拉姆理论被业内人士简称为弱电统一理论，它的核心问题是从根本上解释 W 粒子和 Z 粒子的质量是怎样产生的。按照理论物理学家们的想法，问题的解决需要引入一种叫 Higgs 粒子的新粒子，它是用英国理论物理学家彼得·西格斯（Peter Higgs）的名字命名的。然而没有人确切地知道在哪里能找到这一难以捉摸的家伙，甚至连它们到底是什么样子也不清楚。尽管有许多人仍在寻找这种粒子，但没有人敢断定已经找到它了。

其实在 QCD 理论中，Higgs 粒子基本不发挥什么作用。它实际上是另一套理论——弱力理论中的角色。在这个理论中，W 粒子和 Z 粒子是故事的主角，而电弱统一理论是人们普遍接受的表述方式。例如，弱力会通过一种叫做 β 衰变（beta-decay）的放射性过程表现出来。在 β 衰变中，原子核自发地通过发射电子（在

① 译者注：他们因此获得了 1979 年诺贝尔物理学奖。

放射性衰变过程中称作 β 粒子）和另一种叫做反中微子（anti-neutrino）的粒子而使自身发生改变。W 粒子和 Z 粒子是弱相互作用力（以下简称弱力）的传递者，它们扮演的角色类似于光子在电磁学中担当的角色。

除了在放射性现象和粒子加速器中发挥重要作用之外，弱相互作用力就可以看做是一种宇宙中可有可无的东西了。如果你被放逐到一个荒岛上，不得不抛弃四种力中的一种，你肯定会把弱力丢掉。大自然为什么要不辞辛苦，特地制造一种只能用来破坏一些原子核的力呢？

或许这已经不仅仅是一个人类能考虑的问题了。但在放射性 β 衰变中出现的反中微子，和它的同伴中微子（neutrino），有可能会为我们提供一些线索。中微子（通常用 ν 表示）和反中微子（通常用 $\bar{\nu}$ 表示）是非常特殊的粒子。它们在某些方面很像电子，并且和电子一样，都是真正的基本粒子。但是它们不带电，并且质量极小。它们在强力、弱力和电磁力之中，只参与弱相互作用。最奇特的是，中微子全部都是左旋的，类似地，反中微子全部都是右旋的。

这就意味着，在自然界的所有相互作用力当中，只有弱力能够分辨出左和右。由于弱力为自然界提供了一种区分左右方向的方式，并且还操纵着神秘的中微子，所以实际上，它也许并不像乍看上去那样无足轻重。现在，问题就变成了：大自然为什么要区分左右呢？

初识强力

与之相反，大自然中另一种力（一种物理学家们称之为强力的相互作用力）的存在原因却显而易见。一个典型的原子核包含许多质子和中子，中子和质子的数目要看被讨论的原子核而定。

质子带一个单位的正电荷，相互之间挨得特别近。中子尽管不带电，但也挤在同一个狭小的空间内。是什么让这些粒子①聚集在一起的？无论让它们聚在一起的是什么作用，该作用都必须足够强，以至于在质子间强大的电磁排斥力的作用下，质子还能够牢牢地结合在一起，否则原子核在瞬间就会分崩离析。

而完成这个任务的相互作用就是强力。也正是这种力控制了原子核的“燃烧”过程——核聚变（nuclear fusion）过程，太阳和其他恒星的能量都是由核聚变提供的。在核反应堆内部不断发生着原子核碎裂成更小的核，并且释放出能量的核裂变（nuclear fission）过程，其中强力也发挥着重要作用。在一种叫做 α 衰变（alpha-decay）的放射性现象中，原子核碎裂后会放出一块叫做 α 粒子（由 2 个质子和 2 个中子构成）的“核碎片”，这也是由强力操纵的。

最重要的是，如果没有强力，质子和中子就不会聚在一起而形成原子核。没有原子核就意味着不可能有原子，没有原子我们就不可能存在，不过那样的话也就不会有我们在这里操心这些问题了。

近些年来，物理学家们继续发展并修正了他们曾认为是正确的强相互作用理论。改进后的理论就是 QCD 理论——一套有关强力的理论，这也是本书的主题。

质子和中子本身都是由更基本的粒子组成，QCD 就是写给这些更基本的演员们的剧本。而强相互作用的舞台，也不会比这些演员们所在的世界——原子核更广阔。亚原子世界里，被物理学家们称为“经典规则”的物理规律不再适用，那些经典规则只能应用于尺寸大到可以握在手里的那些物体。所以，为了描述和理解原子内部的世界，物理学家们运用了一套完全不同的规则，即

① 译者注：指质子和中子。

量子力学（quantum mechanics）。他们也逐渐了解到，有关作用模式（pattern）和对称性（symmetry）的研究将是他们继续前进的方向。对实验物理学家们来说，他们正在寻找他们所研究过的最小的东西，为此他们需要最强大的显微镜。那些“显微镜”就是粒子加速器。

第二章 对称性

物理学家们喜欢对称性，这很合乎情理。如果没有对称性，要想弄明白亚原子世界如何运作，是一件十分困难的事情。到目前为止，一切有关基本相互作用力的理论中，对称性自始至终都是理论的基石和支柱。如果不坚定地运用对称性思想，物理学家们在理解弱力和强力时很可能会举步维艰。

数学家们也很喜欢对称性。早在粒子物理学家们发觉他们的工作一刻也离不开对称性之前，数学家们就发现了对称性问题。这种有关对称性的数学被称作群论（group theory），它的起源还要从 19 世纪的挪威数学家尼尔斯·阿贝尔（Niels Abel）和 5 次方程谈起。所谓 5 次方程，是指所含未知数的最高次数是 5 的方程。阿贝尔不断尝试理解 5 次方程的性质，并最终证明了，不存在给出 5 次方程解的简单方法，这一结论引起了和他同一时代的法国人埃瓦利斯特·伽罗华（Evariste Calois）的关注。伽罗华是最具悲剧性色彩的数学家之一，而阿贝尔则一生贫困潦倒。阿贝尔直到去世后两天，才获得一所大学为他提供的职位，他的生平是一部凄惨的传奇。相比之下，伽罗华的生活经历更加丰富多彩，甚至可以用浪漫来形容。他一生饱受政治迫害，还曾经入狱，最终死于一场（与感情纠葛有关的）决斗当中。但就在进行

决斗的前一天晚上，他写了一封长信给他最信任的密友，就是在这封信中，他在阿贝尔计算结果的基础上，构建了现代群论的基础。“群（group）”这个名词也是伽罗华创造的。

有关群论的另一条发展主线与之截然不同，却更能符合物理学家的胃口。它始于法国科学家奥古斯特·布拉维（Auguste Bravais）为晶体对称性进行分类的工作。法国数学家卡米尔·约当（Camille Jordan）把这两种思路和其他一些想法统一到一起，到 19 世纪末，他的工作已经成效显著，群论受到了整个数学界的极大关注。群论能够势如破竹地进军近代物理，还要归功于美籍匈牙利裔物理学家尤金·魏格纳（Eugene Wigner），他在 1926 年研究原子内部的物理时引入了群论。

群论是一个庞大的系统：例如 20 世纪 80 年代完成了对所谓“有限单群（finite simple group）”的分类工程，这项工程耗费了许多数学家多年的时间，记录相关成果的文献累计达 1 万多页。群论还是一个美妙而活跃的领域，它的内容涉及许多其他的数学分支。这一章的主题，就是介绍粒子物理学家们经常用到的那部分群论内容。

描述“相同性”的数学

数学家和物理学家们眼中的对称性，同其他人眼中的一样，都是优美动人的。有时，某个整体的几个部分，从某种形式看上去是完全一样的，这种直觉性的想法激发了人们对对称性的灵感。人的身体关于中轴线大约是对称的，身体的右半边在平面镜中所成的像与左半边（近似）完全相同。早期的群论工作者们对雪花情有独钟，这反映了他们当时的研究兴趣所在。雪花是一种十分迷人的冰的晶体，把它放在显微镜下，可以观察到围绕中心轴排列着 6 个相同的部分。把雪花绕中心轴旋转 1/6 个圆周——

也就是说旋转 60°之后，它的样子和旋转之前一模一样。最后我们再看一个有关圆的例子，圆的对称性同我们在 QCD 理论中要接触到的对称性更为接近。绕着圆心将圆旋转任意角度（无论是几度，还是几分之一度），它看上去都不会有什么变化。

一些变换可以使某些东西看起来并没有发生改变——人们无意中观察到的这类现象，其实包含了一个极端重要的思想。对称意味着存在一类操作，这种操作作用于系统后，系统的状态同初态完全一样。相应地，在系统中也存在一类性质，它们在受到作用后并不发生改变。这些不发生改变的性质，用这种操作的不变量进行描述：对称性就是指相应的不变性——某些性质不发生改变。在实际的物理应用中，不变量的存在，意味着某一物理量的守恒，就像平时在力学中的能量守恒和动量守恒。守恒性质是物理学家们理解和描述世界的核心方式。守恒量所遵从的规律叫做守恒定律。如果没有守恒定律，物理学家们几乎对什么问题都无能为力。而这些至关重要的守恒量都根植于对称性。关于时间变化的对称性对应着能量守恒，沿直线方向位置变化时的对称性对应着动量守恒，而转动下的对称性对应着角动量（angular momentum）守恒。角动量是一种特殊的动量，它的大小取决于转动的速率和转动物体的质量分布。QCD 理论中的对称性与一种新的、奇特的、描述强力的守恒定律有关。群论是 QCD 理论的两个主要部分之一。

正方形的游戏

把一个正方形旋转了 90°后，没有人能看出你是否动过这个正方形。你还可以再这样旋转几次，而无论你旋转多少次，总是不会有人发现正方形有什么变化。正方形关于过其中心的轴具有四重对称性（fourfold symmetry），与之相应存在着一系列操作：转动 90°、转动 180°、转动 270°以及“什么事情都不做”的操

作——也就是说转动 0°或者说根本不做转动。

首先，我们谈谈“什么事情都不做”的操作。数学家们把这种什么都不做的操作叫做“单位元素（identity）”。在上面的例子中，单位元素指的是转动 0°的操作，而在其他情况下就很有可能是把某个数乘以 1 的运算了。其次，你可能会发现，在转动正方形的例子中存在四种转动操作（分别是转动 0°、90°、180°和 270°），对正方形连续做两次这类转动，其效果仍等同于这四种转动操作中的一种①，这种性质叫做封闭性（closure）。换句话说，转动一张平放在桌面上的正方形纸，接着再做一次转动，其最终结果是纸仍在桌面上旋转而不会出现别的情况。第三点是一个技术细节，如果要连续做三次转动，将其分别标记为转动一、转动二和转动三，先做转动一和转动二，再在这两次转动的“合效果”上进行转动三的操作，其结果同先做转动一，然后再进行转动二和转动三的“合操作”的结果完全一样。尽管这两种方法中三次转动的整体顺序都是按照转动一、转动二、转动三进行的，但分组情况不同。这个性质叫做结合律（associativity）。普通数的乘法就符合结合律，但数学家们还知道许多不遵从结合律的例子。第四，存在一个能抵消所做操作的“逆操作”——能使正方形又转回 0°的转动。下面就是一个逆操作的例子：对正方形的每一个转动，都有一个作用相反的转动（四种转动中的一种，如转动 90°和转动 270°），它们共同的作用结果就是单位元素。

因为正方形的转动满足这四个简单的规则，数学家们就称：它们构成了一个“群”。实际上，任何满足这四种性质（有单位元素，符合封闭性，满足结合律，每个元素都有相应的逆元素）的一组元素都构成一个群，这就是群的定义。这些内容虽然看上去平淡无奇，但数学家们能够从十分简单的规律出发，得出许多

① 译者注：如将正方形先转动 180°再转动 270°，等同于直接把正方形转动 90°。

有用的结论。这些规则简单得一目了然，但它们足够描述对称性的本质了。满足这些规律并不是理所应当的事情，数学家们知道许多或多或少违反这些规律的例子。

尽管正方形在平面内的这四种转动构成了一个群，但还存在其他对正方形变换的可能方式，另外还存在四种操作。其中两种是绕正方形的每条对角线各自转 180°，另两种是绕过两组对边中点的直线旋转。因此，能完整地描述一个正方形对称性的群包含八个元素。在正方形所在平面内，绕中心轴旋转的四种操作本身也能构成一个群，这种“包含在另一个群中的群”被数学家们称作一个更大群的子群（subgroup）。

此外，由于每种对称性操作都是将正方形由一种状态跳跃性地变到了另一种状态，因此这两个群都是离散群（discrete group）。这恰好同关于圆的例子相反，对于圆来说，无论做几次任意连续的转动，其状态都和原来一样。与圆的转动相对应的群称作连续群（continuous group）；用来定义群的那一套规则对于连续群也同样适用。连续群中有无限多个元素——旋转 0°到 360°之间的任意角度就是一个元素，每个元素可以用其对应的旋转角度来表示。离散与连续之间的不同，可以用时钟钟面来形象地说明：钟面上 3 点、6 点、9 点和 12 点几个位置是离散的，而秒针转动时平滑地扫过的位置是连续的。

离散群与连续群在物理上都是有用武之地的。例如，在研究晶体结构时，你就会发现离散群发挥了很大的作用。在晶体中，原子排列成规则的点阵，它们对于转动特定角度（例如 90°）具有对称性，对于沿某一方向平移一定步长也具有对称性，这些分立性的对称都用离散对称群（discrete symmetry group）来表述。然而，连续群在物理学中通常更重要一些。原因很容易理解：许多真实的物理系统，都涉及连续的转动、平动以及其他一些物理量的连续变化，这些都与连续性的对称相关。在所有的连续群当

中，一种被称作“李群（Lie groups）”的群对粒子物理学家们最为重要。李群是根据另一位挪威数学家——索菲斯·李（Sophus Lie）的名字命名的。李群本质上是连续群，其涉及的量不仅连续变化，而且是平滑地变化的。这或许涵盖了一个典型的物理学家所能考虑到的所有连续群。在建立有关基本作用力的理论时，碰到的所有的群都是李群。

圆绕其中心轴的旋转所构成的群，就是李群的一个实例。实际上，粒子物理学中所用到的群理论，同有关转动的群理论如出一辙。无论在 QCD 理论中，还是在用以解释质子和中子等基本结构的对称性方法中，我们遇到的都是这种群。

从正方形到圆

秒针的针尖在扫过钟面时，总会同钟面中心保持固定的距离。钟面本身是某个平面的一部分；很容易想象，钟的机械装置都安装在一个平面上，这样秒针的尖端就会在这个平面内划出一个圆形轨迹。一个平面本身可作为一大类变换（指转动变换和平动变换）的背景，这些变换可以将任何一个位置上的物体，移动到其他任一位置上去。与此相比，时钟秒针的平滑圆周运动受到了更多的限制。实际上，秒针尖端的运动是一类受限制变换的例子，叫做正交变换（orthogonal transformation）。“正交”意味着“成直角”——正交变换是指在移动物体或者某个点的过程中，令物体或点与某个固定点的距离保持不变。如果把一个简单的正交变换作用于平面内的一个点上，就会使这个点出现在圆周上的其他位置，因此正交变换就形成了一个连续群，也就是一个关于圆周的连续旋转对称群。

二维平面上的正交变换形成的连续群称作 O（2）群。字母

"O"① 表示"正交"，在有些场合也被称作"单位大小（size one)"的变换。数字"2"是维度数，说明转动是在二维平面内进行的。如果以桌面为参考，一次变换将一个点从它的起始位置移动到一个新的位置上，该点在桌面上沿横向运动的距离可能比沿纵向移动的距离小，但起点与终点都与某个固定点保持相同的距离。换句话说，该变换将 2 个位置坐标值（沿桌面横向的坐标值和纵向的坐标值）变为 2 个新的坐标值。从数学角度来讲，每一个新坐标值都依赖于 2 个原始坐标值，所以将初始值变换到末态值要用到 4 个数。这些数可以写成一个两行两列的阵列，用以表示两个原始坐标值转变为末态坐标值的方式。这些数字所组成的阵列叫做矩阵（matrix)。在数学中，矩阵一般可以包含任意的行数和列数。O（2）群中的"2"来自于转动所在的空间维数，因此也是有实际意义的最小矩阵（2×2 矩阵）的大小。

实际上，尽管矩阵中有 4 个数，但它们是互相关联的，因此这 4 个数完全可以用一个数来表示。这与平面内转动的情况相符，描述这种转动实际上仅需要一个参数——末位置相对于初始位置所旋转的角度。对于最普通的变换——桌面上一点到其他任一点的变换来说，2×2 矩阵由 4 个相互独立的数组成，对应着 4 个独立的参量。

如果将桌面相对于房间倾斜一个微小的角度，就会引进一些新的变化。要解决的还是同样的问题：在桌面上，保持一个点到定点的距离为定长，将该点从一个位置移动到另一个位置的变换应该怎样描述。但我们面临着新的问题，就是我们要用相对于房间的长、宽和高三个方向的值来描述点的位置。将长、宽、高三个坐标值变换成三个新坐标值要用到一个 3×3 的矩阵。尽管矩阵中有 9 个项，描述点的移动实际上仍然只需要一个参数。

---

① 译者注：orthogonal 的缩写。

这看起来像是在用一个更复杂的办法做完全相同的事情，而事实上正是这样。前面的 2×2 矩阵和 3×3 矩阵都对应着同一个群——O（2）群。这两种矩阵只是在具体问题中实现O（2)群的两种不同方式，因此它们叫做 O（2）群的表示（representation)。也就是说，数学家们认识到存在着一个叫做O（2）群的抽象数学实体，并且引入了用 O（2）群来解决问题的思想，比如移动一个点，就意味着能写出这个抽象实体的特定表示。因此，用平面坐标表示——即 2×2 矩阵表示就足够了。要实现涉及空间坐标的相同变换则需要同一个群——O（2）群的 3×3 矩阵表示。群表示所关注的对象可能是桌面上的点，或者是满足某种对称性的基本粒子。一个特定的群会有无穷多个可能的群表示，但其中许多表示并不是相互独立的。

同许多现实中的物理系统一样，时钟秒针有一个性质是O（2)群没有体现出来的。秒针从不会从罩在钟面上的玻璃罩里跑出来，比方说从 3 点处直接跳到 9 点的地方。仅仅要求秒针的针尖距中心固定点的距离恒定，并没有限制这种可能性。而在真实物理系统的情境中，这表示物体可以从一个位置瞬时跳跃到另一个位置，因此出现这种跳变是令人头痛的事情。为此，物理学家们宁可抛弃所有出现跳变的可能。结果得到了O（2)群的一个子群，或者说是一个包含在 O（2）群中的群——SO（2）群。字母“S”表示特殊（special)。如果用矩阵表示群（指将一个点移动到另一点所对应的群)，“特殊性”就体现为：矩阵中数字前的正负号要满足整体的附加条件。并且，所有在基础理论发展中起重要作用的群，都是这种“特殊的”、不发生跳变的群。

这种想法的一个常见推广，是将在桌面上移动的一个点扩展到在整个房间内自由运动的情形。同样，我们还是让这个点到某个固定点的距离保持不变，就能得到 SO（3）群，SO（2)群实际上是它的一个子群。此外，SO（3）群有 3 个相互独立的可能旋

转角，分别沿着空间中上下、前后和左右 3 个方向。所以尽管 3×3矩阵有 9 个项，但实际上仅用 3 个变换角就可以加以表示了。

SO（3）群本身就是物理学中最重要的一种群，此外，它还引入了在粒子物理中十分重要的两种群，其中一种构成了 QCD 理论的基石。

SO（3）群之所以如此重要，是由于它体现了真实转动系统的对称性，例如陀螺、回转仪和旋转的芭蕾舞演员。这 3 种情况都满足角动量守恒。正因如此，陀螺和回转仪才能指向固定的方向，芭蕾舞演员才能通过调整双臂改变他（或她）的质量分布，从而调节转动的快慢。从数学角度上看，角动量守恒对应于 SO（3）变换下系统的守恒方程。

限制在平面内的转动通过一种简单的方式进行组合：对于任意两次转动，无论谁先进行，都会得到相同的结果。换句话说，顺序并不重要。具有这种性质的群称作阿贝尔群（Abelian），是根据尼尔斯·阿贝尔的名字命名的，SO（2）群就是阿贝尔群的一个特例。另外一个常见的例子是电磁对称群（symmetry group of electromagnetism）。

对具有 3 个独立角度参数的转动，问题就更复杂一些了，转动进行的顺序开始起作用：顺序不同，结果也会不同。一种观测这种与先后顺序有关的转动的方法，是将骰子绕两个不同的对称轴转动 90°（图 2.1）。先绕轴 1 旋转再绕轴 2 旋转的结果同先绕轴 2 旋转再绕轴 1 旋转的结果不同。变换进行的顺序不同，产生的结果也不相同，满足这种性质的群叫做非阿贝尔群（non-Abelian）。构成 QCD 理论基础的对称群就属于非阿贝尔群。

到目前为止，我们所考虑的群，都可以通过转动角度赋予明确的意义。对于许多群来说，它们的实际意义并不是一目了然。例如，SO（4）群是四维空间的转动群，它所实现的空间除了具有前后、左右和上下三个长度之外，还要有某个额外的长度，它

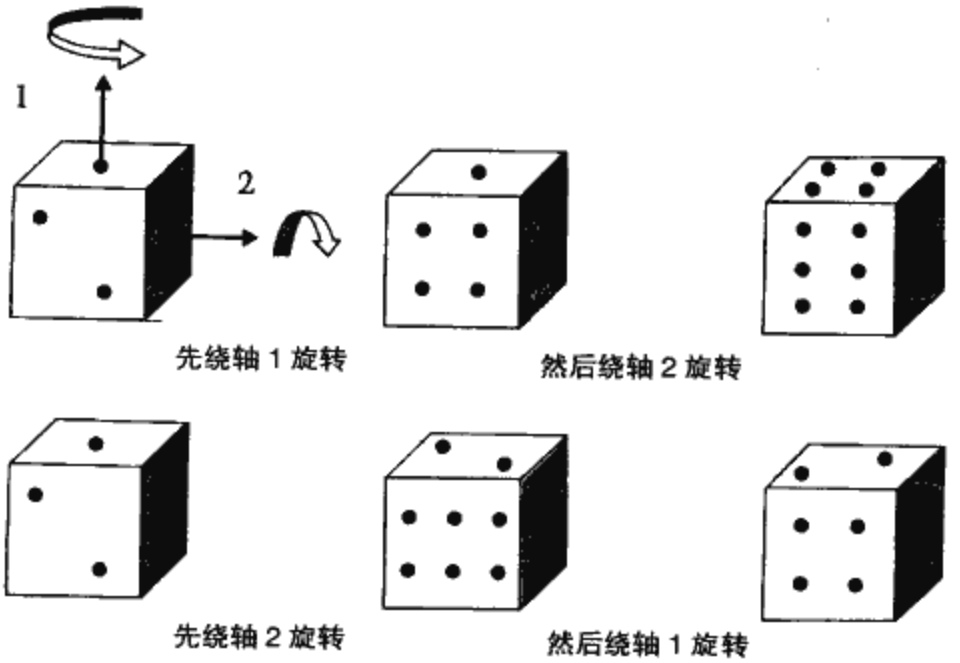


图 2.1　骰子绕两个不同的对称轴转动，两次转动的顺序不同，结果就不一样。在这个例子中，两次不同次序旋转的结果，可以通过进一步绕沿骰子角“伸出”纸面的轴旋转来相互转换

的最小表示是一个 4×4 的矩阵。尽管只有 4 个可能的方向，四维空间中的转动一共需要 6 个旋转角度来表示。因此，人们在头脑中更倾向于把它看做有关几个独立参数的变换，而不是有关角度的变换。或许只有在那些很容易直接验证的简单情况下，用角度描述才成为人们能够接受的方式。在更复杂的例子里，我们可能要处理某个抽象空间中的转动，或者仅仅是“几个参数的变换”。

以上这些想法将带领我们走向何方？下面的例子可能会带给我们一些启示：电磁对称群只有一个独立参数，与其对应的物理事实是：只存在一种类型的光子。相应地，也只存在一个守恒量——电荷。

虽然 SO（2）群和 SO（3）群源自于经典力学中对转动的研究，但它们有着强大的功能，引人入胜。尽管它们与基本相互作用

力理论中的对称性极其相似，并且有着千丝万缕的联系，但它们并不是我们要理解的那种对称性。在有关粒子相互作用的基本理论中，要描述理论的基础框架，就要用到比桌面或房间更为抽象的空间。这类空间的基础与我们平常接触的空间有所不同，譬如在这类空间中，桌面上的两个独立的方向（横向和纵向）有可能会以一种特殊的方式衔接在一起。在这种更为抽象的空间内，等距正交变换被相应的等距单元变换所代替。

粒子物理中的对称性

与平面内的简单转动具有同样效果的群是 U（1）群——也称作一维酉群（unitary group）。这种群与 SO（2）群在数学上是等价的。U（1）群前面没有标注“S”是因为它已经是“特殊的”了，所以加上“S”是多余的。但在其他情况下，确实会有在酉群前面标注“S”的情形，这也是为了排除产生跳变的可能性。SO（3）旋转群本质上就可以用同样是非阿贝尔群、并且也有 3 个独立参数的“特殊酉群”——SU（2）群来替代。这对理解质子和中子间的相互联系以及电子自旋的严格描述都具有十分重大的意义。SU（2）群有一个大姐，叫做 SU（3）群，它也是非阿贝尔群，但它拥有 8 个参数。U（1）群是与量子电动力学（QED）相关的对称群。它与 SU（2）群一起共同说明了弱相互作用的对称性。SU（3）群则描述了强相互作用的对称性，SU（3）群是量子色动力学（QCD）理论中要用到的群。

根据 QCD 理论，强相互作用中最基本的粒子，携带着 3 种标签中的任意一种或几种。这些粒子之间的强力，与 3 种标签的组合方式无关，物理学家们发现这种基本的对称性可以用SU（3）群来描述。在 SU（3）群的作用下，理论中的基本粒子所携带的标签会混合到一起并重新组合。但 QCD 的“主方程”——在本质上支撑着整个理论的方程——在这种群的作用下始终保持

不变。

只要知道与理论有关的对称群，物理学家们就可以利用数学家们已经得到的一切关于群的知识。SU（3）群有 8 个参数，这就意味着在 QCD 理论中要有 8 种携带强力的粒子。就像 QED 领域中，只有光子携带电磁相互作用一样。

以上，我们简短地介绍了对称性怎样进入了 QCD 领域。除了对称性之外，QCD 理论的另一个主要部分，是一套可应用到微观世界的力学体系：量子力学（quantum mechanics）。

第三章 量子世界

有史以来最优秀的理论

在磁场中，电子的行为像小磁铁。它们有南极和北极，当把电子放置在离磁铁很近的地方时，电子就会受到一个使它们扭转的力，使这些小“电子磁铁”沿着磁场的形状排列。这种扭转力的大小，由电子的一种内禀性质——它们的磁矩（magnetic moment）来决定。

电子（就这个问题而言，可以简单地把它看做一块条形磁铁）磁矩的大小，等于扭转力的大小同引起扭转力的磁场强度之比。换言之，在磁场给定的情况下，磁体的磁矩越大，所受的扭转力就越大。但实验结果发现，电子的磁矩与理论最初所预言的不一致。其修正因子被称作反常磁矩（anomalous magnetic moment）——这是一个无量纲的物理量，用符号 a 表示。实验物理学家们多次测量过这个值，得到的结果精确得难以置信。1987年，华盛顿大学（位于西雅图）的罗伯特·范·戴克（Robert Van Dyck）、保罗·施温伯格（Paul Schwinberg）和汉斯·德默

尔特（Hans Dehmelt）用限制在彭宁势阱中的单个电子测量到的值为：

$a_{expt}=0.001\ 159\ 652\ 1884\pm0.000\ 000\ 000\ 0043$。

我们可以对比一下康奈尔大学的木下藤一郎（Toichiro Kinoshita）在1999年得出的理论值：

$a_{theory}=0.001\ 159\ 652\ 1535\pm0.000\ 000\ 000\ 0029$，

木下藤一郎几十年都致力于计算和比较 a 值的工作。“a”值究竟是多少其实并不重要，重要的是理论值和实验值竟如此接近。它们的一致程度精确到了8位有效数字，当理论预言和实验测量相吻合的时候，我们就得到了最好的理论。QED理论常常能得到这种精确的结果。正是这种一致性验证了我们的推断：QED是有史以来最优秀的理论。

QED（量子电动力学）理论名称中的“量子”，源自于这种理论可以直接应用于原子尺度的世界。“电”说明理论是与电磁场有关的，“动力学”是指理论描述了场随时间变化而变化的规律。QED理论就是一套讲述像电子这样的带电粒子，怎样在最小尺度上发生相互作用的理论。用美国物理学家理查德·费曼（Richard Feynman）的话说，QED理论就是“关于光和物质的奇特理论”。

QED理论是我们已知的、有关物质通过电磁力如何进行相互作用的最基本描述。由于所有的原子和分子，都是通过电磁力相互作用聚在一起的，因此它们的行为也加深了我们对电磁相互作用的理解。并且，由于包括我们自己在内的所有物体，都是由原子和分子组成的，QED理论在某种意义上，解释了一切化学规律以及其他范畴内的许多现象。

尽管电子磁矩是一种对QED理论的严格检验方式，实验物理学家们还是仔细地用其他方法来检测理论预言的结果。例如，电子有一个表亲——一种叫做 μ 介子的基本粒子，它也具有反常

磁矩。从 1996 年开始，实验物理学家们就在纽约布鲁克海文国家实验室，用一台巨大的专用仪器观测数以亿计的 μ 介子衰变。他们测量到的 μ 介子反常磁矩值，直到小数点后 6 位，与理论值完全一致。

气体中的放电现象，可以使气体发出光和无线电波，这些辐射具有特定的复杂光谱。在历史上，意义最为重大的例子是氢气的光谱。氢光谱含有一个特殊的高频辐射贡献，即所谓的兰姆位移（Lamb shift），经典理论对此无法解释。但按照 QED 理论，这是很自然的事情。实验测量值与 QED 理论的计算值吻合到 5 位有效数字。另外一些实验测量了氢原子的其他谱线以及其他一些原子的谱线，所有的测量结果都与 QED 理论的计算值高度一致。

在大型正负电子对撞机（the Large Electron Positron collider，简称 LEP）中，电子同它的反粒子——正电子发生碰撞，同时产生 2～3 个光子。CERN 的 L3 实验小组用巨大的粒子探测器研究了这个过程。与光谱和反常磁矩实验相比，该小组以及全球各地的粒子加速器实验室，在更高的能量范围内证实了 QED 理论的预言。

研究固态物质的物理学家们在一个完全不同的环境下，为 QED 理论提供着进一步的支持。凝聚态物理中的两种奇特的效应——量子霍尔效应和约瑟夫逊效应，给出了 QED 理论的强度参数，该参数与通过电子反常磁矩所推导出的值一致，符合精度达到了千万分之一。

这些事实都极有启发意义，但又能说明什么呢？那些走在大街上的男男女女对这些现象一点儿也不了解，它们的重要性又体现在哪儿呢？之所以会出现这种情况，是因为，尽管 QED 理论被称作适用于万物的理论，但要细致地研究它，会遇到许多抽象的问题。反常磁矩的测量以及其他一些实验，都为 QED 理论提

供了精确的验证。游戏的目的是要找到一些仅由 QED 理论就能解释的过程，过程中所测量到的量要与计算值精确吻合。即使过程中涉及其他效应，这些效应也是可估算的，只涉及 QED 理论的部分可以分离出来。从一大块煤的性质中或者其他生化反应中，寻找验证 QED 理论的可靠实验是没有意义的，尽管 QED 理论是它们的基础，但这些系统太过复杂和混乱。因此，只能用一些人们不熟悉的现象作为理论的验证。

还有另一个问题亟待解决，那就是：为什么要讲这么多有关 QED 理论的内容？这些内容与 QCD 理论有什么关系？QED 理论经过反复的验证与检验，最终得到人们的认可。它作为经典电磁学发展的产物，也继承了经典电磁学完美无瑕的基因。QED 理论是一个（相对来说）简单而优美的理论，是 QCD 理论的模板。QCD 理论同 QED 理论十分相似，有共同的基础，但 QED 理论细节上没有那么麻烦。这些细节不久就会讲到，之后还会谈到 QCD 理论的所有重要性质。但我们还是要先讲一些基本的东西。

量子力学

量子力学中的“量子”同量子色动力学（QCD）中的“量子”是一样的。量子力学描述的是小尺度世界（一般不会比一个原子大）中的行为。由于强力的作用范围是原子核，有关强力的任何理论必须是一套符合量子力学的理论。因此，要真正地理解 QCD 理论——有关强相互作用的理论，就要了解量子力学所支配的奇特的微观世界。

法国东南部的格勒诺布尔（Grenoble）古时候是一座要塞城市；现如今，它拥有一座世界上独一无二的研究机构：马克思·冯·劳厄-保罗·郎之万（Max von Laue-Paul Langevin）研究所（简称 ILL），就坐落在风景迤逦的阿尔卑斯山脚下，它是世界上

进行中子研究的主要研究中心之一。

ILL 并不研究中子本身的性质，而是利用中子来进行研究。在 ILL 以及世界上其他一些类似的研究所中，科学家们利用中子的特殊性质来解决各种各样的问题——从铁轨承受的压力到光合作用，无所不包。ILL 的设备主要依靠核反应，在仪器中，铀核裂变产生大量中子，中子通过管道被输送到许多实验区当中。ILL 的反应堆是世界上最强的中子源，并且安装的附属设备也最多。

中子是一种粒子，具有很小但是可以测量的质量，而且还具有其他一些明显的“像粒子一样”的性质，这些性质都可以测量和逐一列出。

当一个中子进入某种物质内部时，由于强力，它会被原子中心的原子核反弹出去。这同电子与 X 射线散射的情况恰恰相反——X 射线是被原子中的电子散射的。中子散射为其他的散射方法提供了必要的补充，因此是一种十分有用的实验工具。中子散射具有很奇特的性质：被散射的中子作为粒子，被环绕在样品四周不同角度处的探测器所记录，最终得到的中子密度分布图样却是波动所特有的干涉图样——中子是一种波。

两列同时传播的波，波峰与波谷并不需要严格地同步，一列波的波峰比另一列提前或延后的的量叫做相位差。如果在波的传播过程中，相位差保持不变，那么就称这两列波相干。相位本身体现了波的周期演变，就像辐条与车轴的夹角表征了车轮的转动一样。实际上，相位是一个角度，它的取值范围对应着一个圆周内的角度值，也就是说它可以取 0°～360°的任意值。波的相位部分对波的振幅（也就是波的强弱）没有贡献，因此，相位具有圆周的几何性质。

相位和相干性是理解干涉现象的基础，而干涉现象又为物理学上的许多精确测量技术提供了依据，为物理测量带来了一个黄

金时代。例如，路易斯安那州和华盛顿州的引力波探测器就是利用了激光探测技术，在4千米长的探测器臂上，测量出只有不到原子核直径的千分之一那么长的长度变化。

如果两列相干的波相遇，例如照到屏上的两列精心制备的光波，一起到达的两个波峰会发生叠加，振幅变为原来的2倍，亮度会变为原来的4倍。波峰和波谷相遇则会相互抵消，产生一个低强度的区域。这就是干涉。两列叠加的波是加强还是削弱，是由不同的相位决定的。

追溯到艾萨克·牛顿的时代，按照牛顿的想法，光可以看做是一种叫做“光微粒”（corpuscle）的粒子。这种想法被另一个英国人——托马斯·杨的工作推翻了。杨在1802年做了物理学上最著名的一个实验：让一束光照射到两条平行的狭缝上，他本来天真地认为，在狭缝后的屏上应该看到两条亮线——狭缝的投影，但他却在屏上发现了一组明暗相间的条纹。当托马斯认识到这是干涉现象后，他断定光一定是一种波。干涉图样是由从两条狭缝中射出的光会合后产生的：波峰与波峰叠加使光增强，形成较亮的区域；而波峰与波谷相互抵消形成较暗的区域。

两个世纪过去了，科学家们仍然在做着与杨氏双缝实验类似的实验。奥地利物理学家安东·策林格（Anton Zeilinger）和他的同事1988年在ILL做了中子的杨氏双缝实验。他们把硼丝夹在用硼玻璃制成的夹子上，让中子束射向硼丝，硼丝和夹子之间的两个空隙就形成了两道狭缝。如图3.1所示，他们的结果十分清楚地揭示了“中子波”的干涉图样，因此他们声称，他们的实验是当时对“物质波”存在的最精确的验证。在ILL工作的无数研究者们，都在利用中子的波动性质做实验，但这还不够，还必须用经典的证明波动性的实验来证实中子的确是波。

此外，人们还做了整个原子的双缝实验。第一次完成这个实验的是德国康斯坦茨大学的奥利弗·卡纳尔（Oliver Carnal）和

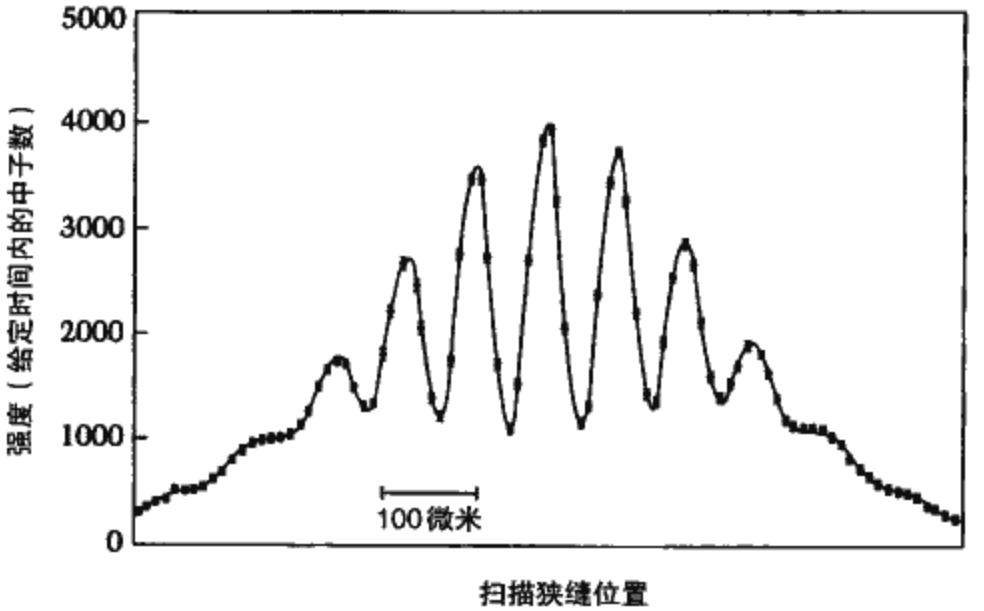


图 3.1　通过双缝的中子产生的干涉图样，中子计数的峰和谷能让人们回想起经过双缝的光形成的明暗相间条纹。实线表示计算值。图片来源：安东·茨林格/《现代物理评论》

约根·米利内克（Jürgen Mlynek），在 1991 年的实验当中，他们让一束氦原子流射向刻在金箔上的两条 1 微米宽的狭缝，在狭缝后观测到了原子波的干涉。在 1995 年，美国麻省理工学院（MIT）的普利查德（Pritchard）研究小组观测到了钠的双原子分子干涉现象。

现在，问题有点严重了。原子和中子一直被认为是构筑物质世界的基本砖块，这就意味着它们一定都是粒子吗？这些实验中却出现了干涉现象：原子和中子都具备波动性行为！

那么中子、原子、电子以及其他一些物质，到底是波还是粒子呢？或许这个问题最好的答案就是：问题本身就是不正确的！实验显示，大自然具有内在的波粒二象性。根据操作或观测方式不同，中子、原子及其他物质会表现出波动性或粒子性。但这两种性质不会同时表现出来。波动性和粒子性都是物质的内在性

质，适用于宇宙中所有的物质。

用来解释波粒二象性的理论体系就是量子力学。尽管波粒二象性是普适的，但在原子或更小的尺度上，它的作用更加显著。因此，量子力学有时也被称作亚微观世界的力学。在亚微观世界里，日常物体所遵从的力学规律就不再适用了。

量子世界

量子力学中的“量子”源自于对光的认识。人们发现，曾经被托马斯·杨证明是波动的光，只能以一些小份能量的整数倍（即能量子）的形式存在。给定光的频率，光量子的能量就是频率和一个常数（称作普朗克常数）的乘积。光量子通常被称为“光子”（photon）。由于每个光子的能量太小了，光子的数目又十分巨大，牛顿很难为他的粒子说观点找到直接的证据，但他的光微粒说最终还是取得了胜利。

普朗克常数是以马克思·普朗克（Max Planck）的名字命名的，正是普朗克发现了世界的量子本质。1900 年，普朗克引入了“能量子”（quantum）的概念，用以解决长期悬而未决的光辐射和光吸收的问题。5 年之后，爱因斯坦又引入了光量子的概念。普朗克常数用 h 表示，它的值特别小，只有 6.625×10^{-34}（单位是焦耳·秒，记作 Js）。现在，普朗克常数已成为最基本的宇宙常数之一。但物理学家们会更常用到 $\hbar$，它的值等于 h 除以 2π，$\hbar$ 在方程中出现的次数特别多。

普朗克常数的值太小了，导致自然界中物质的波动性在日常情况下显现不出来。一个“高尔夫球波”的波长大约是 6×10^{-34} 米，这么小的波长不可能对击球产生任何影响。公路上行驶的一辆汽车波长更小，只有约 2×10^{-38} 米。这些数字都小得不可思议，比原子核的大小还要小许多个数量级，与之相比，光波（本来很小）的波长也显得十分巨大了。绿光的波长大约是 5×10^{-7}

米，家里用的电视机中电子的波长约为 7×10^{-12} 米，所以即使是电子的波长也是很小的。

由于这些波长都太小了，高尔夫球员和司机们永远都不会直接感受到量子世界的奇特。由于量子高尔夫球或量子公路可能会给人们带来许多不可思议的现象，所以这很可能是一件好事。双缝实验显示的不仅仅是“粒子”的波动本质，还提供了更多的线索。实际上，双缝实验可能揭示了量子效应最根本的性质。

如果用光做双缝实验，当挡住其中的一条狭缝时，干涉条纹就会消失。这是因为干涉条纹需要从两条路径通过的光重新会合。在双缝实验中，两条路径分别是通过两条狭缝的两条路线。如果我们将光源变暗，光源射出的光束将会减弱，直到每次只有一个光子射到缝上，这样会发生什么事情呢？日常经验告诉我们，光子要么会穿过其中一条狭缝，要么会穿过另一条。但日常经验是不正确的：单个光子实际上同时通过了两条狭缝。一个光子经过两条狭缝的共同作用后在屏上形成了一个小点。当有足够多的单个光子通过双缝后，屏上最终会形成干涉条纹——由一个个单独的点所组成的疏密相间的图样。

从 1927 年开始，芝加哥大学的亚瑟·登普斯特（Arthur Dempster）和哈罗德·巴索（Harold Batho）进行了更加精巧和令人信服的实验。他们降低了光源的强度，证明了单个光子可以同自身发生干涉。一些持怀疑态度的人可能认为，他们可以通过观察光子到底从哪条狭缝中通过来解决这个问题。但他们如果这样做的话就会发现，观测行为本身会破坏干涉条纹。1982 年，美国物理学家首次提出了一种“量子擦除器”（quantum eraser）。1992 年，加州大学伯克利分校的保罗·科威特（Paul Kwait）、埃弗雷姆·施坦格（Aephraim Steinberg）和雷蒙德·乔（Raymond Chiao）完善了这一装置，该实验漂亮地演示了上述现象。

图 3.2 是双缝量子擦除实验的示意图。每次只有一个光子到

达双缝。通过上方路径的光会经过一个“标识器”，它通过某种方式标记光从上方的路径经过。两条路径会汇聚到一块分光镜上。分光镜是一块部分镀银的镜子，能反射50%的入射光，而让另外的50%通过。分光镜的功能只是为了提供一个让两条路径汇合、叠加的场所。在分光镜与屏间还有一台用来消除“标识器”所作标记的装置，这个消除标记的装置就是“量子橡皮（擦除器）”：路径信息被存储在光子的偏振态中，可以用偏振仪去掉，其具体细节就不在这里详述了。

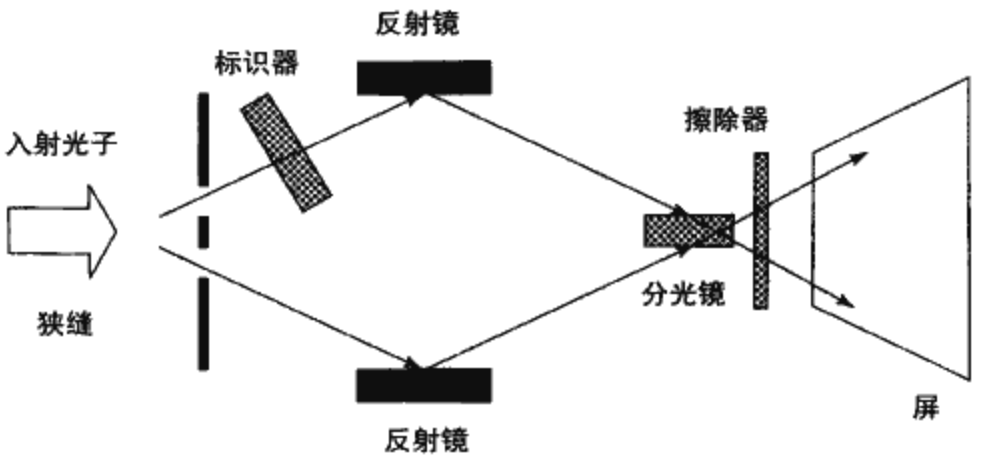


图3.2 双缝量子擦除实验的示意图。箭头是入射光子可以经过的路径。两条路径在分光镜上会合。标记和擦除分别在两种叫偏光器和偏振仪的装置上实现

如果没有标识器和擦除器，这就是一个简单的单光子杨氏双缝干涉实验，它会出现干涉图样。如果只有标识器而没有擦除器，干涉条纹就会消失。如果既存在标识器，又存在擦除器，干涉条纹将会再次形成。

之所以这样，是因为如果只有标识器，从上方经过的光子会带有路径信息，也就是说明它从哪条路径经过，或者通过哪条狭缝的信息。一旦光子经过的路径可以识别，干涉条纹就会消失。

加上擦除器之后，尽管从上方经过的光被做了标记，但在到达最终的探测器之前，标记就已经被清除了，两条路径不再可以分辨，干涉图样就会再次产生。

只存在标识器而不存在擦除器时，光子的行为像是粒子，每次只通过其中一条狭缝，并且不会产生干涉。加上擦除器之后，光子的行为就像波，会同时经过两条狭缝。量子世界唯一的秘密就在于，光在到达双缝时怎么会“知道”之后是否会遇到擦除器，从而来决定表现为粒子性或是波动性？

把电子、光子以及其他一些东西看做是真正的粒子或波，都是毫无根据的。最好的解释是，电子有时显示出粒子性，而有时显示出波动性。或者用威廉姆·布拉格（William Brag）先生的话说就是：“物理学家们在周一、周三和周五运用波动理论，在周二、周四和周六运用粒子理论。”对粒子物理学家们来说，他们习惯性地把任何东西都称作粒子（这从他们的称号里也可以看出来），尽管他们自己心里也知道，这是不完全正确的。

几率与不确定性

“粒子”波动性的一个重要结果就是，它的精确位置和动量不能被同时确定。如果把一个电子看做是波，那么谁还能说清楚这个电子究竟在什么位置？粒子的位置和动量都有内在的不确定性，两者满足基本的量子不确定性关系，这是由粒子的波动性所引起的：位置不确定度与动量不确定度乘积的最小值等于普朗克常数。

这就是海森伯的不确定性原理，由沃纳·海森伯（Werner Heisenberg）在 1927 年发现。不确定性原理，是对自然界内在的不可测量性的描述，与实验的具体技巧没有关系。它阐述了，即使在理论上，也不可能找到一种方法，来同时测量位置和动量的精确值。它也说明了，要提高某一个量的测量精确度，势必会导

致另一个量不确定度的增加，两个量的不确定度之积是一定的。

在粒子物理学中，不确定性关系会不时地以另一种等价的形式出现。在这种形式中，粒子能量的不确定度和测量持续时间的不确定度，同样遵守海森伯的不确定性原理。如果能量值已经确定，粒子具有该能量的时间就会完全不确定，反之亦然。不确定性关系可以用来解释各种各样的现象，例如宇宙怎样从虚无中诞生。

波粒二象性没有回答一个“电子波”或光波怎样变成了被探测器接收的几个“电子”或“光子”。双缝实验为此提供了一些线索。

第一条线索是，将双缝干涉图样中最亮的部分和单缝图样中最亮的部分作比较。双缝图样中的亮度不是单缝图样中的 2 倍，而是 4 倍。波的大小（振幅）变成原来的 2 倍，强度就会变成 4 倍，这说明强度是振幅的平方。第二条线索来自于这样一个事实：如果将双缝中的一条关闭，特殊的干涉条纹将会消失。双缝图样的强度，并不只是一条缝单独存在时产生的图样和另一条缝单独存在时产生的图样强度的叠加，这种特殊的图案只能由从两条缝通过的波的振幅叠加形成，叠加后的总振幅的平方就是图样的强度。

其规律就是这样：先将振幅叠加，然后将总振幅取平方，得到了最终的强度，这就是波动怎样起作用的。但这只是答案的一部分。一种波动形式怎样才能变成屏上或某种电子探测器上所得到的可数的粒子呢？

每个“粒子”经过双缝到达探测屏形成一个单独的点，最终许多点所组成的图案将会形成干涉图样。一个单独的点对应着一个单独入射的电子或光子，它自己不能产生特殊的密度图样。但在最终很亮的区域里，找到一个点的几率会很大，而在最终显得比较暗的区域中找到一个点的几率就会很小。

是几率！几率（或称作可能性）是宇宙中不可缺少的组成部分，更是量子力学中必不可缺的一部分。由于具有波动性，你永远都不可能准确地知道有着确定动量的电子或者光子此时此刻在哪儿，或是一会儿将会出现在哪儿。但你却能给出它们出现在某个位置的可能性有多大。量子力学也无法提供一套计算原子中的电子会跃迁到哪个动量的方法，它只能给出发生某种跃迁的可能性。

在量子力学理论中，要把所有有贡献的量子波幅先叠加起来，然后再取平方算出某种可能性。实验者只能在测量的时候观测到最后的可能性，波幅本身则被隐藏了。

薛定谔与氢

物理学家们从量子态出发，讨论量子理论。例如，某个量子态可能是一个电子某个时刻位于空间内的某一个点。第二个量子态可能是同一个电子在稍后的某个时刻位于空间内另外的一个点。量子态也可以包含其他的属性，其中一些会在相关的课程中讲到。

根据量子理论，某个事件发生是通过系统从一个量子态跃迁到另一个量子态完成的。一个简单的例子就是一个电子从一个地方移动到另一个地方，也就是从某时刻在某点处的量子态跃迁到稍后某时刻另一点的量子态。量子力学提供了计算这种跃迁振幅的规律。把这些振幅取平方就得到了可能性，这样就可以同实验比较了。

描述量子态变化最简单的方式就是薛定谔方程（Schrödinger equation），是由埃尔温·薛定谔（Erwin Schrödinger）在 1926 年提出的。其最经典的应用就是解释了氢原子的光谱。薛定谔方程描述了原子中心的正电荷对单个电子量子态的影响。这样，可以计算出一系列不同的电子“能级”，每个能级都具有确定的能量。氢原子

中的电子可以通过吸收具有适当能量的光子，从能量较低的能级沿着能级阶梯爬到能量较高的能级上面去。相反的过程也同样成立，电子可以从能量较高的能级跌到能量较低的能级上去，同时释放出光子，光子的能量恰好是两个能级之间的能量差。薛定谔方程甚至还给出了计算这种跃迁概率的准则。

从历史上讲，将量子力学运用于氢原子，可以给出电子跃迁的能级，并且伴随着光子的吸收或放出，这是具有重大意义的。这不仅是因为它解释了氢原子为何能稳定存在（该问题在非量子的电动力学中是无法解释的），而且它很自然地解释了氢原子的光谱现象。光谱是一系列精确的、具有特定频率的光，表现为不同频率的光具有不同的颜色，氢光谱可以通过氢气的放电过程产生。在量子理论诞生之前，科学家们无法解释这种现象。理解氢气以及其他气体的发射光谱和吸收光谱（气体会吸收经过该气体的特定频率的光，从而形成吸收光谱），在量子理论发展的各个阶段都起着举足轻重的作用。

薛定谔方程构成了量子力学中许多应用的基础。例如，化学家和原子物理学家们都会时常用到它。运用薛定谔方程的最简单的例子，是把原子核仅仅看做静电荷源，电子与原子核之间满足库仑静电定律。库仑定律说的是，当两个静电荷之间的距离增加为原来的 2 倍，它们之间的静电相互作用力会减少为原来的 1/4。这是被物理学家们称作“平方反比定律”的一个例子。同样，静电相互作用能与静电荷之间的距离成反比。所以，尽管电子本身遵从量子规律，它同原子核之间的相互作用却是用老式的、非量子的（或称作经典的）静电规律来处理的。

这种处理方式可以得到许多漂亮的结果，但这种做法在理论上并不完备，并且还存在许多其他的问题需要用一套完整的量子理论来解决。薛定谔方程仅仅是一个开始。

最短路径

检验员用光线来校直建筑物的墙壁或道路两边的台阶。他们之所以这样做，是基于两个重要的假设：光沿直线传播，并且光只能沿同一条路线前进。那些检验员们很幸运，因为房子和马路都很大，波动和量子效应并不明显。双缝实验说明，在研究亚原子尺度的问题时，这两个假设都不完全正确。

如果起点和终点之间存在两条狭缝，一个电子或光子将会从两条路径上通过。没有任何限制规定我们只能设置两条狭缝，因此我们可以设置上千条、上万条狭缝，那么同样就会存在成千上万条对结果有影响的可能路径。为什么要费这么多力气来设置如此多的狭缝呢？如果我们把狭缝之间的格档全部去掉，就可以等效地认为在连接起点与终点之间的干涉空间内存在着无数条路径了。

在双缝的情形中，最终出现的密度图样是由两条相关路径的相位差控制的。在存在数以亿计的可能路径时，情况也会是这样。对于那些远离“合理路径”的路径，它们的相位各不相同，各条路径对结果的贡献相互抵消，因此离得比较远的路径贡献很小。对于离“合理路径”比较近的那些路径，情况就有所不同了，这些路径的贡献会大得多。各条路径所提供的无数个小振幅的和，构成了从初态跃迁到末态的最终量子振幅。

这种方法叫做路径积分。在 20 世纪 40 年代早期，理查德·费曼（Richard Feynman）还在普林斯顿读研究生的时候就提出了这个方法。这是最强大的量子力学公式，薛定谔方程也可以由路径积分推演出来。薛定谔方程对于化学家和原子物理学家们的工作来说，已经足够了，但对于一套描述基本相互作用力的理论来说，几乎没有能替代费曼的路径积分的方法。

路径积分计算了连接两个量子态之间无穷多可能路径的所有

贡献，连那些弯得特别厉害的路径也包括在内。每条路径对最终结果的贡献都依赖于它的相位，最终结果是由理论的“主方程”所决定的。

路径积分理论中最奇妙的一点是：所用到的主方程是一个简单的经典非量子方程。把一个经典的主方程放入路径积分理论中，就会把原来的理论变成量子理论。这是因为一旦置身于路径积分理论中，主方程就控制着来自无数条路径的所有贡献，而不仅仅是经典力学中的一条路径。路径积分理论用到经典的主方程，说明经典理论和量子理论的一致性是显而易见的。最终，理论所需的许多附加性质都可以从主方程引进。从根本上讲，可以通过路径积分理论自动得出最终的量子理论。在研究基本相互作用理论的奇特对称性时，路径积分理论本身也足以胜任。

用路径积分的方法，理论物理学家们利用对称性和其他一些特性，推出合适的主方程，以此为起点构筑起量子理论。然后，他们可以把主方程放入路径积分理论中，根据一些现有的方法可以得到一系列计算法则。接下来，他们就可以利用这些计算法则，算出物理上可测量的量，譬如不稳定粒子衰变的快慢，或者是粒子碰撞时的情况。用路径积分的方法，很容易看出主方程的变化对计算法则的影响。

在路径积分中所用到的主方程并不神秘：它们表现了系统的总动能与相互作用能之间的差别。例如主方程可以是电子动能与光子动能之和，减去电子和光子之间的相互作用能。把不同的主方程代入到路径积分理论中，会得到不同的量子理论。

可以用一个骑自行车的例子，来帮助我们熟悉一下基本的经典思想。你在起点和终点之间有几条路线可以选择。最短的路线需要你先直接爬到山顶，然后再下山，到达山的另一边。从山顶骑着自行车下坡是将重力势能转化为动能的过程，这当然很舒服，但蹬着自行车上山会把人累得半死。还有一条路线是穿过一

条十分平坦的山谷，但需要多走几千米的路程。你最终选择了第三条路线：沿着山脚绕过去。这样既不会像穿越山谷那样沉闷乏味地走很长的路，也不会像爬过山顶那样累得筋疲力尽，而是在痛苦与无聊之间选择了一个折衷的方案。你选择了一条最佳路线，但从地理意义上来说，这并不是一条最好的。因为它既不是一条最短路线，也不是最平坦的，这是考虑了道路是否平坦、路程长短以及劳累程度等各个方面的因素才做的决定。

自然界也会选择最佳路线。大自然不知道什么是痛苦，什么是无聊，但它却关心能量问题。当骑自行车的人在对着地形图选择路线时，大自然也在精明地打着算盘。在一个过程中，运动能（常称作动能）和势能（指与位置有关的能量或者是相互作用能）的总量是保持不变的。过程中保持动能不变需要花很长时间，因而是不可取的，这就像穿过山谷的那条路线一样。经过势能极值点的路线也是不可取的，就像经过山顶的那条路线。大自然最终所选择的是在过程中，动能与势能之差对时间的积分取最小值的那条路线。能量的交换值叫做“作用量”（action）。自然界选择的最佳路线就是作用量最小的那条路线，这就是最小作用量原理。

作用量是力学的核心内容。数学家们知道如何才能够用简单的方法找到作用量的最小值，根据最小作用量原理，就可以得到经典力学的有关定律了。一个在空中飞行的棒球会将它的作用量减小到最小，它会沿一条曲线飞行。最小作用量原理的一个简单推论是：从检验员们所用的激光器中射出的光束，是按照直线传播的；而放置在游泳池底的光源发出的光，射出水面时会发生弯折。在量子理论中，经典的最小作用量原理被路径积分所取代：主方程以作用量的形式被引入，它通过相位控制着每一条可能路径的相对重要性。

在上高中的时候，费曼第一次领略到了最小作用量原理的神

奇之处。他的博士论文也是以“量子力学中的最小作用量原理”为题目的。

当量子力学遇到相对论

相对论是一个会永远改变我们对宇宙的理解的伟大思想，但我们用两句话就足以将它概括出来。第一，在两个惯性参照系（其中一个参照系以某一个恒定的速度相对于另一个运动）中，所做的任何相同的实验，其结果总是相同的。第二，光速在所有参照系中都是同一个常量。

究竟什么是惯性参照系呢？我们稍候再详细说明。不管它是什么，总之，实验者在两个不同的地方做同一实验，都会得到同样的结果，这看起来并没有什么意思。如果每个人测量光速时都会得到同样的结果，你是否觉得会稍微有意思一些呢？

这些看似平常的陈述中，实际隐藏着极其丰富的内容。它们是爱因斯坦提出狭义相对论（special relativity）的基础，也是广义相对论（general relativity）诞生的前驱。广义相对论描述了引力作用，它除了研究匀速运动的系统外，还研究有加速度存在的系统。

狭义相对论之所以在有关 QCD 理论的故事中占有一席之地，是因为要想让 QCD 理论成为严格的关于强相互作用的理论，就必须让它满足狭义相对论所提出的规则。其次，由于所涉及的粒子能量和速度极大，高能物理的许多表达方式都是相对论的直接结果。最终，狭义相对论成为了我们理解许多粒子重要性质的基础。

在一根静止的导线附近摇动磁铁，就可以使导线中产生电流。或者让磁铁保持不动，但晃动导线，导线中同样会产生电流。

爱因斯坦在 1905 年那篇著名的论文，就是以对摇动磁铁和导线这个问题的简单讨论开篇的，从而把他的狭义相对论公布于世。它对于磁铁和导线问题的观点是：只有相对运动才有意义。无论磁铁和导线是运动还是静止，只要它们之间存在相对运动，导线中就会产生电流。所有的惯性参照系对于所有的物理实验都是等价的——正是这个观点，引导着爱因斯坦迈向了相对论原理。

惯性参照系的意义实际上没有什么特别之处。一个参照系就是一套可供测量的坐标轴。实验者可以通过参照坐标轴来观测实验，例如观测一个物体的运动。如果没有参照系，就没有办法判断物体是否正在运动。惯性参照系是一种简单的参照系，在惯性参照系中，自由运动的物体相对于参照系没有加速度。一个固定在地球上的实验室并不是一个惯性参照系，因为实验室中的每件东西由于地球引力的作用，都有一个向下的加速度。在太空中漂浮的宇宙飞船是惯性参照系的一个经典例子。

在这种意义上，惯性参照系有一点被特殊化了。实际上，只有从理论上解释相对性时才会用到它。这也是为什么一讨论相对论，就经常要谈到观测者 O 和“嗖嗖”乱窜的宇宙飞船 O′。其实，惯性参照系就是一个没有引力或加速度干扰的实验室，实验室中的物体都保持其本来的状态。假设有两个这样的实验室，它们之间以恒定的速度相对运动，两个实验室中的实验者都会观测到同样的物理规律。这就是爱因斯坦的相对性原理的内容。

这就意味着不存在进行实验的特殊地点，在太空中也不存在普遍意义上静止的巨大实验室。没有特殊的实验地点，在哪里测得的实验结果和物理定律都可以作为其他实验的基准——这是一个意义十分深远、基本而又简单的思想。

第二个定律——光速对于所有参照系都相同——也是由它的电磁起源决定的。电磁场的麦克斯韦方程组，揭示了一个电荷静

止的时候仅仅是一个电场源。当电荷运动时，它同时也成为一个磁场源。如果在观测的时候引进相对性：可以让电荷保持静止而实验者运动，实验者所观测到的电磁效应应该不发生改变。

麦克斯韦工作最重要的成果，是预言了电磁扰动会以光速在真空中以波的形式传播。人们第一次发现的电磁波是以无线电波的形式出现的。电磁波包含一套完整的波谱，频率范围从无线电波、微波、可见光、X射线一直到γ射线。它们在真空中的速度完全相同，都等于光速，光速可以用基本的电磁学常数表示出来。由于没有理由让我们认为这些常数会随做实验的人不同而不同，这就说明光速对于所有的人来说同样也是常数。也就是说，光速总是不变的，是一个大得不可思议的数：3×10^8 米/秒，这将会导致一些奇怪的结果。

设想一个静止的足球运动员，朝着另一个倒霉的迎面而来的运动员把球踢过去。忽略摩擦和空气阻力，第二个运动员将会受到足球的冲击，冲击的速度等于球离开静止运动员的脚时的速度加上第二个运动员的奔跑速度。这种简单的速度相加的办法，就是我们平时处理这类问题的方法。

但是换一种情形，让一个穿白衣服的静止科学家，用手电筒照向朝她跑过去的一个同事，跑步的科学家可以测量从手电筒发出的光的速度。测量结果与她们都静止时的结果完全一样。第二个科学家测量到的光速值同书本上给出的一模一样，也就是光束离开手电筒的速度。根据日常经验，这是一个十分荒谬的结果。

但这却是世界的本来面貌，实验结果也正是这样。例如我们不让那个穿白衣服的科学家打手电筒，而是让一种叫做中性π子（或者叫π^0介子）的粒子衰变为两个光子。如果所用的π^0介子的运动速度与光速大小相近，实验者会观测到，朝π^0介子运动方向发射的光子仍然以光速传播，而不是光速与π^0介子的速度之和。同样，朝与π^0介子运动相反方向发射的光子也会以光速运

动，而不会由于 π^0 介子的运动使它的速度减慢。我们平时所用的力学对于这些现象已经无法解释了。

综合这两条规律——在不同的惯性参照系中所做的实验总会得到相同的结果，并且光速恒为常量——爱因斯坦发展了一套新的力学。这些惊人的现象只有在速度非常高时才会变得明显。

两个实验者可以用光信号对比结果，这为相对论的理解提供了帮助。表述两个坐标系中长度和时间之间联系的变换叫做洛仑兹变换（Lorentz transformation），这是以荷兰物理学家洛伦兹（Hendrik Lorentz）的名字命名的，他在爱因斯坦提出相对论的几年前，就推导出了这些公式。如果一个实验者有一把 1 米长的尺子，那么当另一个实验者经过他身旁时就会发现，这把尺子会比 1 米要稍微短一些，具体的值可以通过洛伦兹变换式计算出来。相应地，第一个实验者也可以看第二个实验者的尺子，他也会发现尺子缩短了。这就是相对论思想的一个推论。没有一个参照系会比其他参照系更特殊，因此，在两个实验者之间总会存在令人费解的对称性现象。这种效应叫做长度缩短效应，我们很难将它同日常经验联系并统一起来。在加利福尼亚的斯坦福线性加速器中心（简称为 SLAC），原长为 3 千米的线性加速器可以把电子加速到十分接近光速的极高速度。在这种速度下，长度缩短效应十分明显，3 千米长的加速器对于如此之快的电子来说，只有大约 1 米长。

时钟也会有类似的效应。如果两个实验者都有钟，并且也可以看到对方的钟，他们都会发现对方的钟跑得比自己的慢。这种时钟变慢的现象被称作时间膨胀效应，这在粒子物理实验中是很普通的现象。一些衰变粒子的寿命比它本身的要长，就是因为它们在实验中运动得特别快。

洛仑兹变换也给出了在高速情况下求合速度的规则，但并不是像日常生活中用到的那种简单相加的方法（像此前足球运动员

的例子中的那样）。用新法则求两个速度的合速度，所得的结果永远都不会超过光速。

这些相对论世界的奇特性质，在粒子物理学中比在日常生活中更加重要。因为在平时，物体运动得都很缓慢，只有当速度接近光速时，相对论效应才会变得显著。日常生活所用到的力学，实际上是相对论在低速时的近似。

时空弯曲

一只满屋子“嗡嗡”乱飞的苍蝇，是一个运动缓慢的、非相对论性的普通物体。要描述苍蝇在屋子中的位置需要 3 个数：第一个数用来表示苍蝇在沿屋子长度方向上的位置，另一个用来表示它在沿屋子宽度方向上的位置，第三个数表示它距离地面的高度。换句话说，房间内部是一个三维的空间。

过了一段时间之后，假设苍蝇已经移动了一段距离，它的位置又可以用 3 个另外的数给出。时间是每次测量之间的间隔，在日常生活中，它是与空间（如长、宽和高）完全分开的。爱因斯坦第一个认识到时间和空间并不是相互独立的，时空是以某种简单的方式相互纠缠的。只有当两个以极高速度相对运动的实验者，相互对比他们的时钟和尺子时，这种纠缠效应才会显得十分明显。

苍蝇从这儿飞到那儿所经历的时间间隔是多少？在非相对论情况下的实验者们，很容易在这个问题上达到一致。他们也会对苍蝇从刚才的位置飞到现在所在的位置，所经过的距离达成一致。他们所认同的是由“距离定则”所给出的总距离，而不是他们按照各自的参照系所测量到的那 3 个表示空间位置的数。距离定则的内容是：先求 3 个方向上的距离的平方和，然后再取平方根，所得的值就等于空间内两点之间的距离。这有一点像直角三

角形中的毕达哥拉斯定理①。这中间没有任何技巧，如果你找一把尺子直接量一下两点间的距离，你得到的测量结果会和距离定则给出的结果一致。

在狭义相对论中，实验者们只会认同一个修改过的、由时空间隔复合而成的“距离”，可以通过一种清晰的方法把它计算出来。人们不能在通常意义上的空间距离和时间间隔上达成一致，因为它们各自都要遵从洛伦兹变换的表达式。实际上，洛仑兹变换是一套完备的变换形式，通过它可以推导出这种“修改后的距离”表达式。原则上，一套能保持某些量完备性的变换可以构成一个群，洛仑兹变换所构成的就是洛伦兹群（Lorentz Group）。

按照经验，也是狭义相对论中比较重要的一点，就是在包含极高速度的情形下，用四维时空的方式处理问题，会比用三维空间加上第四个与众不同的时间维度更加严谨。

让我们再次回到那间有苍蝇的房间。连接苍蝇刚才所在的位置和现在的位置之间的线段，也可以用 3 个数来表示。如果这只苍蝇想要逃跑，那么它在长和宽方向上移动的距离值可能会很小，但高度变化却会很大。这条线段也具有一个方向，就是从苍蝇刚才所在的位置指向现在的位置。就像一个提示怎样追踪苍蝇的箭头一样。一条由一组数和一个方向所确定的线段叫做向量（vector，也叫矢量）。而像苍蝇质量这种只有大小没有方向的物理量，被称作标量（scalar）。实际上，由于三维向量历史最久且最为常用，人们常常把三维向量简称为向量，而把前面的定语“三维”给省略掉了。

在狭义相对论中，要运用两个实验者都认同的“修改后的距离”，就必须用到由四个数和一个方向所确定的向量，这就是四维向量。它是用来讨论狭义相对论和粒子物理问题的一种基本物

① 译者注：即勾股定理。

理量。四维向量就是满足相对性变换（例如洛仑兹变换）的四个数，其中三个数表示某一事件的空间坐标，另一个表示时间坐标。为了证明这一点，我们可以计算四维向量（而不是三维向量）的改变量，其大小正好等于时间间隔同光速的乘积。以极高速度相对运动的两个实验者，可以把 4 对（8 个）数放入“修改后的距离”公式，求出他们所认同的两个事件之间的“时空距离”。除了连接两个点之间的向量——位置向量之外，另一类重要的向量是速度。我们还要再看一下刚才那只苍蝇，这有助于了解苍蝇飞得多快以及它的飞行方向，以便打苍蝇的人能够迅速完成任务。速度的大小与方向构成了一个速度向量，由于只需要三个数就可以确定苍蝇的速度，因此它也是一个三维向量。

如果发生了苍蝇和另一个飞行中的物体相撞的事情，那么用另外一类向量——动量来处理问题会比用速度更加有用。从数学上来讲，动量是质量与速度的乘积；从物理上来讲，动量描述了朝某个方向运动的某个东西有多重以及它运动的快慢。当苍蝇撞到别的东西或者粒子发生相互作用时，守恒（即保持不变的）量是动量而不是速度。因此，在描述物理过程的时候，动量总是比速度常用得多。

在高能物理中当然也是这样。粒子在探测器中以接近光速的速度交错前行。狭义相对论预言，最简洁的方式是利用从普通动量扩展而来的物理量——四维动量，这是另一个四维向量的例子。四维动量在具有很高速度的粒子对撞中保持守恒，这使它成为物理学家们讨论和测量粒子相互作用的常用语言。

为什么要这么麻烦地使用四维向量呢？这是因为四维向量的基础是所有人公认的“修改后的距离”定则，用四维向量表示的任何关系对每个人都适用。方程的形式在洛仑兹变换前后保持不变，这就意味着如果某个实验者得出了一个物理定律，并把它用四维向量的形式表示出来，那么这个方程将会对其他所有的实验

者都成立。这个相当方便的性质叫做形式不变性，或协变性(covariance)。

尽管四维动量之间的关系是恒定不变的，但四维动量的测量值却依赖于进行测量的观测者。但有一些量在所有的参照系下都严格保持同一个值而不发生改变。例如，电子的静止质量对于所有参照系都是同一个值。“修改后的距离”定则可以把两个四维位置向量复合起来；只要用同样方式复合两个四维向量，都会得到一个无论谁来观测都不会改变的数。如果一个四维向量按照这种方法同自己复合，就得到了所谓的“四维向量的平方”，这是其中的一个特殊情形。粒子四维动量的平方总是等于粒子静止质量的平方。它们都是洛仑兹不变量——在洛仑兹变换前后保持不变。

为了让这些内容不显得太抽象，我们有必要说明一下，粒子加速器的能量就是用这种方式表达的。例如，位于 CERN 的质子-反质子对撞机，也称作超级质子-反质子同步加速器（Super proton-antiproton Synchrotron，缩写为 $Sp\bar{p}S$），其能量可以达到 630 个能量单位（关于“能量单位”的具体含义，我们稍候再详细介绍），W 粒子和 Z 粒子就是在这里首次发现的。但是，用“加速器的能量”这样的字眼来讨论问题可能会有误导之嫌，因为一束粒子流如果迎面撞上另一束与之类似的粒子流，就会比它与静止的“目标粒子”相碰时具有更多的“有用能量”。因此，粒子物理学家们常会用到一个叫做不变质量的量（用 $\sqrt{s}$ 表示），它只与碰撞粒子的四维动量有关，并且，对于所有观测者都具有相同的数值——即具有洛仑兹协变性。实验物理学家们在绘制粒子碰撞的数据图表时，经常把能量 $\sqrt{s}$ 作为关系图的横轴。在 CERN 的质子-反质子对撞机中，两个粒子束碰撞的不变质量是 630 能量单位。但其中一束粒子轰击静止目标时，它的不变质量就会小得多，能够产生粒子的有用能量只有约 25 能量单位（而

不是 315 能量单位）。粒子碰撞实验的结果以及理论物理学家们的努力都指出，那些粒子碰撞的效果必定与进行观测的人无关。它们对于所有参照系必定是相同的。实际上，这仅仅是把所有数据都填到一个表格中加以验证，以确保它们都满足爱因斯坦提出的“物理定律在所有惯性参照系下都相同”的要求。

这对于我们研究有关强相互作用的理论以及有关物质的其他一些基本相互作用理论的意义在于：只有满足洛仑兹协变性的理论才有价值。这是多重检验的第一步，任何不能经受这种检验的理论注定只是一堆垃圾。QCD 理论满足洛仑兹协变性，因此我们可以通过这一关，接下来处理其他问题了。

混为一体的质量与能量

在当今世界上，有一个方程“尽人皆知”，那就是爱因斯坦的质能关系方程 $E=mc^2$。其中，E 代表能量，m 代表质量，c 是真空中的光速。这是我们之前所介绍的那些想法的自然结论，它表述了一个事实：所有能量都具有质量，当然其中也包含了一些令人吃惊的结论。

一个运动的物体会具有动能，也就是在交通事故中弄弯保险杠和破坏其他东西的那种能量。出现在质能关系中的质量表明，除了通常公认的物体“静止质量”（当物体静止的时候测得的它的质量值）之外，还存在与动量相对应的附加质量：当物体运动得越快，就会显得越重，这是另一种在粒子物理学中很常见的现象。实际上，出现在质能方程中的质量 m 并不是物体在静止时的质量，而是运动物体的质量，它会比物体的静止质量大。例如，在电视机里的电子束中的电子质量，会比书中给出的电子质量的值稍微大一些，书中的值只适用于静止电子。而那些绕着粒子加速器“嗖嗖”做圆周运动的电子质量会显得重得多。这也是可以解释为什么任何物体都不能被加速到超过光速限制的一种方式，

因为那样做需要无穷大的能量。

通过质能关系，可以很明显地看出高能物理学家们所用到的特殊能量单位的形式。任何人都有权利使用一套能让自己生活更为便利的单位系统。例如在计算一次旅途所花费的时间时，如果用国际通用的标准单位秒来衡量，将会是一件十分痛苦的事情。旅行社一般都会用小时来标注飞机的航程，这就足以确保飞机的时间并且更加有效。另一方面，秒（或者零点几秒）在奥林匹克运动会的百米赛场上却是很实用的单位。另外，由于任何系统都是相对论性的并且是量子的，光速和普朗克常数会在各种方程中经常见到。

物理学家们为了使生活变得更简单，使他们的方程显得更加明了，规定光速和约化普朗克常数的值都为 1。这样做的效果是使这两个量在所有方程中完全消失。这也意味着粒子的静止质量要用兆电子伏（Mega Electronvolt，一种能量单位，缩写为 MeV）来表示。例如，在这种所谓的“自然单位制”中，质子的质量是 938.3 兆电子伏，换算成常用单位是 1.673×10^{-27} 千克；电子的质量是 0.511 兆电子伏，也就是 9.109×10^{-31} 千克。

那么电子伏是什么东西呢？它是一个能量单位，相当于把一个所带电荷与一个电子电荷相同的粒子，经过 1 伏特电压后加速所得的能量，可以记为 eV。例如，把一个 12 伏的汽车电池，接到两块相距几厘米的金属板上。从接到负极的金属板上产生的电子，会向接正极的金属板加速运动，当它达到正极时，会具有 12 电子伏的动能。电子伏是一个非常小的单位，把 1 升水的温度升高 1 度需要消耗 2.6×10^{21} 电子伏的能量。

在原子中来回运动的电子的特征能量是几个电子伏。核物理学家们会对更高的能量感兴趣，像几千电子伏（千电子伏缩写为 keV）或几兆电子伏（兆电子伏缩写为 MeV）的能量。粒子物理学家们更熟悉上千兆电子伏（吉电子伏，缩写为 GeV）的能量，

最近，他们已把所涉及的能量值提高到了几百万兆电子伏（太电子伏，缩写为 TeV）。现在回到我们前面所举的 CERN 的质子-反质子对撞机的例子中，其中提到了能量单位的问题，那台对撞机实际上具有 630 吉电子伏的能量。

再加上量子力学

当薛定谔提出他的量子力学方程时，狭义相对论已经诞生 20 多年了。尽管它已经取得了巨大的成功，但很明显，薛定谔方程从一开始就与相对论格格不入。要找到一个满足相对论的量子理论是一项艰巨的任务，这个任务是由一位年轻的英国物理学家保罗·狄拉克（Paul Dirac）完成的，他的工作成果斐然。

相对论性的量子理论需要满足洛仑兹变换下的不变性，并且一定也要满足相对论性的质能关系，而且只能给出正的几率。第三个条件有点奇怪：因为最早建立相对论量子理论的物理学家们最直接的做法将导致负几率的出现，负几率是个无法理解的概念。最直接的做法还有另外一个明显的问题，那就是粒子会出现负能量，这也是无法理解的。

1928 年，狄拉克写出了一个满足相对论条件的电子的量子力学方程。他认为在方程中应该平等地考虑时空，而这种概念在薛定谔方程中是不存在的。狄拉克的做法，只有在提出一些全新的数学解释以及给电子赋予一种叫做自旋的属性时才能成立。此外，狄拉克的方程还存在另外一套解，尽管不存在负几率的问题，但这套解表明存在具有负能量的粒子。因此它几乎将狄拉克方程毁于一旦。凭借着非凡的勇气和直觉，狄拉克将负能量问题变成了一次奇迹般的飞跃，他提出，方程的“负能解”描述了一大类新的物质——反粒子。

我们先来看看自旋。狄拉克方程和狭义相对论，并不是电子自旋第一次在物理学中出现的地方。电子具有自旋是由荷兰物理

学家塞缪尔·古德斯密特（Samuel Goudsmit）和乔治·乌伦贝克（George Uhlenbeck）在1925年提出的，他们提出这个概念是为了解释原子光谱中的双线现象——这是原子光谱对量子理论发展起过重大作用的又一见证。自旋也就意味着，带电荷的电子的行为就像一个小磁铁，这是由于运动的电荷是磁场的源头。从运动电子的角度来看，原子核所带的电荷是磁效应的源头。因此，自旋的电子实际上就是磁场中的一个小磁体，它的能量也与它的自旋是沿着磁场方向还是逆着磁场方向有关。这就决定了会存在两条紧邻的能级。而正是这种能级的分裂导致了光谱中出现的双线现象。

钠光谱是一个著名的双线光谱的例子。钠蒸汽中的电火花会放出特殊的黄光，这实际上不是一条谱线，而是两条——钠光谱D线。相应的波长分别是589.6纳米（缩写为nm）和589.0纳米，如果没有电子自旋，就只会存在一条谱线了，其波长大约是上面两个波长的平均值。

狭义相对论显示，自旋是我们所生活的时空的产物，我们所在的时空具有洛仑兹变换群所包含的特殊对称性。洛仑兹变换的实质不过是一些拉伸和旋转，从这里也可以看出这种特殊关系的一些端倪。将其运用到实际中，例如一个电子的运动中，电子的四维动量应该在洛仑兹变换下保持不变，而自旋是最简单的转动，用SU（2）群描述。自旋是根据洛仑兹变换的最自然描述提出的，实际上，狄拉克方程也可以通过这种方法推导出来。

电子自旋，是由宇宙的量子本性所决定的，正是这种量子本性限制了电子自旋的测量值。比如地球绕地轴自转以及其他一些经典力学中的问题，在原则上那些物体可以以任意速率绕任何轴自转。电子不喜欢这种自由。在磁场中放入一个电子，并规定一个参考方向，习惯上称作z方向，我们可以沿这个方向在实验上测量电子的自旋，结果是电子会具有顺时针自旋（电子自旋只是

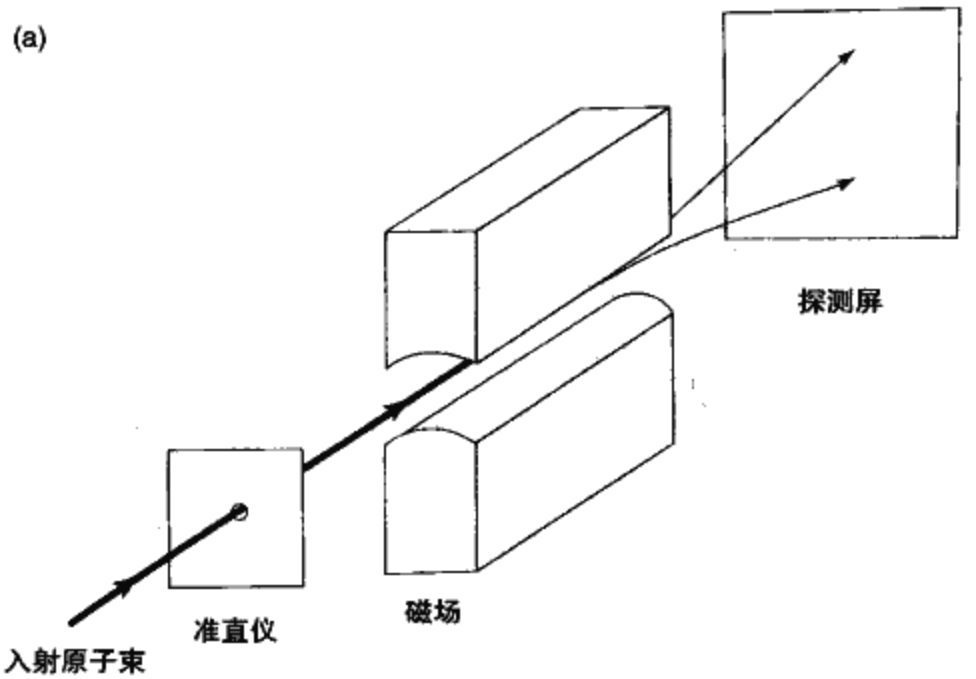
一种字面意义）或者逆时针自旋。并且，其自旋的大小总是固定的，为$\hbar/2$。习惯上经常把$\hbar$省略掉，简单地把电子自旋表示为1/2。电子自旋的测量结果要么是+1/2，表示自旋向上；要么是-1/2，表示自旋向下。还存在另一种表示自旋的有用方法。自旋向上也被称作顺时针或右手系——可以用右手做一个翘大拇指的动作，让拇指指向电子运动的方向，而弯曲的四指表示旋转的方向。自旋向下则可以表示成逆时针或左手系。即只存在两个分立的值，它们分别对应于原子光谱由单线分裂而成的两条线。

电子自旋的强度值为1/2，这是与实验中观测到的电子与其所在的磁场之间相互作用强度相一致的。能级分裂为两条表明有两个自旋态，并说明能级不会取连续的值，而只能取分立值，因此可以说自旋是量子化的。与电子自旋的大小相联系的数字，就是量子数的一个例子，量子数是指量子态的实验测量值。

量子测量定则，不仅限制了沿某个方向的测量结果为特定的值，并且也限制了是否可以同时沿其他方向测量自旋的值。把电子放入一个磁场中，并选取一个作为参考的z方向，然后测量出自旋值为+1/2或-1/2，根据定则，这就意味着不可能同时测定沿x轴方向和沿y轴方向的自旋值。

如果让一束银原子穿过一个经过特殊设计的磁场，就会清楚地显示出客观世界对于电子自旋的限制以及对实验者了解客观世界的限制。银原子的原子核周围环绕着47个电子，其中有46个电子组成了一个围绕原子核的对称球体。第47个电子的自旋决定了银原子的磁矩，因此也控制了银原子在磁场中的行为。

由以上分析可知，当一束银原子穿过一个在垂直于原子束运动方向上有变化的磁场时，如果存在该方向上的磁场梯度，每一个原子都会受到一个与电子自旋大小成正比的力，见图3.3中的图（a）。当磁场确定之后，力只与磁场梯度同电子自旋沿z方向部分（称作自旋的z方向分量）的乘积有关。因此，磁相互作用



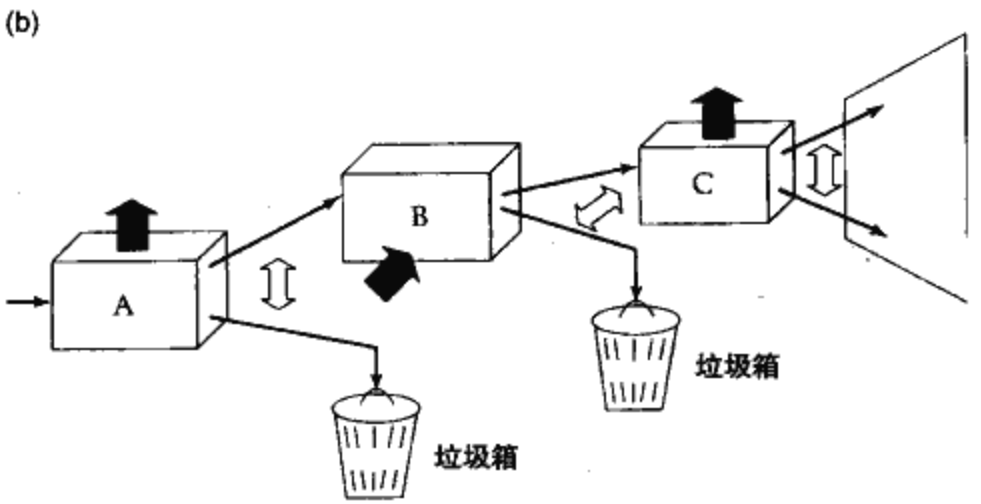


图 3.3 图 (a)：斯特恩-格拉赫实验。一束银原子被磁场梯度垂直分为两部分。图 (b)：由三个这样的实验构成一个实验序列，中间那个的磁场梯度同其他两个实验呈直角。每个盒子代表一套类似于图 (a) 的那种分裂原子束的磁装置

对方向十分敏感，可以挑选出电子自旋平行于（或者反平行于）磁场梯度的原子。

实验者会观测到什么现象呢？如果电子自旋的行为是经典的，那么电子自旋的 z 方向分量可以取某个范围内的任意值。相应地，原子束中不同的原子，在 z 方向上的受力会在某个范围内连续地变化，结果会在探测器上出现一个平滑变化的峰值。这个实验已经有人做过了，最开始是 1922 年奥托·斯特恩（Otto Stern）和沃尔特·格拉赫（Walther Gerlach）做的，但他们所观察到的却不是我们所想的这种现象。在他们的经典实验中用到了一束银原子，在没有磁场的情形下，他们在玻璃板上得到一条银原子线。当他们打开磁场，在他们原来看到一条线的地方，他们观测到了两条原子沉积线。在经典理论中，这条线应该变成一条具有宽度的模糊区域。斯特恩和格拉赫实验表明，磁场将原子束严格地分成了两个单独的部分。他们对实验结果感到十分震惊，于是给哥本哈根的尼尔斯·玻尔寄了一张明信片，玻尔那一年因他在原子结构方面的工作，刚获得诺贝尔奖。他们的明信片上印上了原子在无场情况下的单线图案，旁边还印上了存在磁场情况时的分裂图案。

电子自旋值在 z 方向上的分量，如果以 $\hbar$ 为单位的话，会是简单的 $+1/2$ 或 $-1/2$，而不是经典力学所预言的，可取这两个值之间的任意值。换句话说，这种测量发现了电子只能取自旋向上或者自旋向下两种态，而不存在中间态。在测量之前，某个电子会处于自旋向上和自旋向下的不确定态。测量行为使它变为一种或另一种可能的态。一旦你发现某个粒子处于自旋向上的态，紧接着再做第二次同样的测量时，你就会发现电子还会处于自旋向上的态。

但是，如果第二次所做的是不同的测量，会出现什么情况呢？这可以用 3 次连续的斯特恩-格拉赫实验来说明，见图 3.3 中的图（b），前一次实验的结果可以用作下一次实验的条件。图中，每一个盒子就表示一套图 3.3 的图（a）中所示的仪器。从盒

子 A 中分离出的自旋向下的电子都被扔掉了，自旋向上的电子继续通过盒子 B。盒子 B 中磁场梯度是沿着水平的 y 方向的。从盒子 B 出来后，电子束又被平均地一分为二：其中一束电子的自旋沿着 y 轴方向，而另一束电子的自旋与 y 轴方向相反。自旋与 y 轴方向相反的那一半电子也被丢掉了，剩下的电子继续通过最后一个盒子：盒子 C，盒子 C 是与盒子 A 完全一样的。最终结果是，从盒子 C 中出来的电子在沿着 z 轴的方向上有一半自旋向上，另一半自旋向下。

进入盒子 B 的电子束是 100%自旋向上的；但根据盒子 C 的测量结果，当电子束离开盒子 B 时，却只有 50%的电子是自旋向上的，而另外 50%的电子是自旋向下的。这些自旋向下的电子就像变魔术一样，不知道是从哪里冒出来的。这就说明：尽管盒子 B 能够清楚地测量出电子自旋沿 y 轴方向的情况，但盒子 B 中的这种测量却把电子自旋究竟向上还是向下（指自旋沿 z 轴方向的情况）搞得一团糟。在一个盒子中，同时精确地测量自旋沿 z 轴方向的情况和自旋沿 y 轴方向的情况是不可能的。

换句话说，盒子 B 中所做的“测量”破坏了通过盒子 A 所得到的所有信息（包括进入盒子 B 的电子束以及被抛弃的那部分电子）。同样，随后在盒子 C 中所进行的测量，反过来又破坏了进入盒子 C 的电子束所具有的全部信息。以上分析说明，每一个盒子都会将输入电子的量子态设置为两个可能的态之一。而这两个量子态是由盒子中磁场的方向决定的，与电子束原来的态无关。

要描述银原子在磁场中的行为，只需要一个量子数就够了，这个量子数就是银原子最外层的那个电子相对于某个选定轴测量到的自旋值，它的值只能取+1/2 或−1/2。不存在用来表示沿另一个方向的自旋值的其他量子数。这个性质反映了一种与自旋相联系的基本对称性变换——SU（2）群。SU（2）群具有三个参数，相应的变换中只包括一个维持量子态（量子态用量子数表

示）的参数。正是 SU（2）变换的这个用一个参数表示的子集，对应着一个量子数——这揭示了量子理论和群论之间的深刻联系。

自旋很难用图画出来，因为自旋本质上是一种量子属性。有时认为电子“自旋向上”是一种十分有用的想法，但这种简单的“经典图像”会给人以相当大的误导。实际上在经典力学的框架里，一个“电子球”必须旋转得特别快，结果会导致电子的某一部分运动得比真空中的光速还要快，这是同物理定律相悖的。事实上，没有人看到电子真的在沿顺时针或逆时针自转。

自旋是一个相当重要的属性，除了能造成原子光谱的分裂之外，它还是为粒子进行分类的重要依据。电子以及其他一些粒子具有半整数倍的自旋（自旋值为 1/2，3/2，5/2 等），这类粒子称作费米子（fermion），是根据意大利物理学家恩里克·费米（Enrico Fermi）的名字命名的。那些具有零自旋和整数倍自旋（自旋值为 1，2，3 等）的粒子被称作玻色子（boson），是根据印度物理学家萨地扬德拉·玻色（Satyendra Bose）的名字命名的。通常，自旋为 1 的粒子——像 W 玻色子或 Z 玻色子有 3 个分立的可能自旋态，分别用＋1，0 和－1 来表示，这是量子世界的又一特征。光子尽管十分普通，但在某些方面却是相当特殊的。光子的静止质量为零；正因为如此，它的一个可能的自旋态——被标记为 0 的自旋态是不存在的：尽管光子的自旋为 1，但它只具有两个自旋态。

下面，我们再来讲一下狄拉克方程转败为胜的经历。为了将相对论和量子力学统一起来，狄拉克提出了一些前提条件，这导致了他的方程除了得到一组平常的解之外，还存在另一组与之对称的解。这组解看上去对应着一种与电子一模一样，但具有负能量的粒子。

狄拉克天才地把这组负能解解释为，存在具有负能量的反粒

子（antiparticle）。根据狄拉克的想法，电子会有一个与之具有同样质量、同样自旋但电荷相反的反粒子伙伴。这种物质——正电子（positron）在1933年被人们首次发现，从而证明了狄拉克的伟大预言。

根据狄拉克的理论，反粒子是宇宙的时空性质和量子本质的必然结果，这些性质并没有限制电子反粒子的存在。出现在弱相互作用中的中微子是一个左撇子，它也是一种自旋为1/2的粒子，因此，它也应当存在并且的确存在与之相应的反粒子——反中微子。同样地，质子也有反粒子，叫做反质子。反质子是1955年被发现的，它的质量与质子质量相同，并且具有一个单位的负电荷。反中微子是人们在1959年发现的。

狄拉克于1984年逝世，死后被安葬在佛罗里达州。狄拉克方程取得了巨大的成功，作出了重大的贡献，位于伦敦的威斯敏斯特大教堂（Westminster Abbey）为了纪念狄拉克，在1995年树立了一块牌匾。这块匾与牛顿墓毗邻，上面记录了狄拉克的卓越成就。

量子力学要满足狭义相对论，其结果就是相对论性量子力学的诞生。这也引领我们走向对粒子自旋以及反粒子的理解。相对论性量子力学也揭示了原子光谱的一些极其精细的细节，以及表示自旋与电子磁矩强度之间关系的常数。这个常数的值是2，在非相对论量子力学当中，该常数的值必须“人为规定”；但是在相对论性量子力学中，可以通过狄拉克方程推导出来。

实际上，该常数的实验测量值要比2稍微大一点点，因此我们夸大了狄拉克方程处理该问题的能力。狄拉克方程在解决兰姆位移问题上也存在问题。如果根据狄拉克方程，根本不会出现兰姆位移现象。另外还存在其他更深层次的问题，使狄拉克方程无法成为一个完整描述基本相互作用力的理论。狄拉克方程已经步入正轨，却不足以抵达目的地。

量子力学中的诞生与湮灭

狄拉克的反粒子为那些正在寻找新粒子以及探索已知粒子之间相互作用的物理学家们提供了帮助。例如，在 KEK 实验室（Koh-Ene-Ken，日本国家高能物理实验室，坐落在东京以北的筑波市，距东京约 60 千米）工作的实验者们利用具有 8 吉电子伏能量的电子和 3.5 吉电子伏能量的正电子，在一个 3 千米长的圆形非对称能量对撞机（称作 KEK-B）中发生对撞，在此过程中产生大量的粒子流。在伊利诺伊州巴特维亚市的费米实验室中，质子与反质子在强大的万亿电子伏加速器 Tevatron 中迎面相撞，同样也会产生大量的粒子流。在以上两个例子中，粒子与反粒子会相遇、湮灭；与此同时，其他的粒子与反粒子会在它们湮灭的地方诞生。

“普通”的量子力学曾经被用来描述氢原子能级及其他一些问题，但在处理这些问题时却显得特别棘手。因为这种形式十分简洁的理论，不包括粒子产生和消亡的内容。而这些是很多高能粒子反应的基本性质：两个粒子会互相碰撞，然后发生湮灭，在余烬中又会有其他的粒子产生。

粒子的诞生和毁灭无处不在。例如一个高能伽马射线光子，可能是由地球的大气层中一个正在减速的带电粒子发出的，该带电粒子本身是来自太空的“宇宙射线”的产物，而它又能自发地蜕变为一组电子-正电子对。这个过程叫做电子偶的产生（pair production），是粒子物理中一种寻常的现象，表现了物质吸收伽马射线的一种标准机制。当一个电子和一个正电子相遇时，它们也会互相湮灭，同时在湮灭处产生两个光子，这就是著名的电子偶湮灭（pair annihilate）过程。

即使是最简单的原子跃迁问题也暗藏着隐患。当一个电子吸

收了一个光子，跳到更高的能级上时，光子去了哪里？当一个电子从比较高的能级跌到低能级上时，会发出光子，这个光子又是从何而来？在第一种情况下，光子消失了；在后一种情况下，又会产生光子。但是光子究竟到哪儿去了（或从哪儿来），它又是怎样消失（或产生）的？粗略地讲，这就是探索量子理论的先驱者们最初遇到的粒子产生与湮灭的问题。他们最终解开了这个谜，从而也创建了量子场论。运用量子场论，我们能够理解自然界四种基本相互作用力中的三种，只有引力作用还徘徊在量子场论的门槛之外。最成功的量子场论是 QED 理论，它为另外两种理论——QCD 理论和电弱相互作用理论提供了范式。那么究竟什么是量子场论，它又是怎样解释量子力学中的诞生与湮灭现象的呢？

量子场

“场”是一些分布在某个区域内的量或者影响，而不是仅仅存在于某个点上的东西。例如气象图可以很好地说明这种基本观点。在气象图所表示的整个区域内，温度会平滑而连续地变化——这说明在这个区域存在一个温度场。在理论上，区域内的每一点以及任何一点的温度值都可以给出。同样，气象图上的每一点都会有风，因此同一个区域上会分布着另一个场。在这个场中选定一个点，就要用一个数字表示风速大小，还要用一个箭头表示风的方向。

风和温度形成的场是两个实际中的例子，分别代表了两种不同类型的场。温度只有大小，但没有方向，因此是一个标量，所以温度场是一个标量场。而风既有大小又有方向，因此形成了一个向量场。可以用一个布满小箭头的区域形象地把这种场表示出来，每个小箭头的方向代表该点的风向（向量的方向），而小箭头的长度代表风速（向量的大小）。

场的例子无处不在。例如，地球是一个巨大的条形磁铁，具有明确的南极和北极。用罗盘可以在地球周围探测到其产生的磁场。用来把电能从发电站输送给用户的高架电线会产生电场，为了证明这一点，可以把一个荧光管放到电线的正下方（但千万不要这么做!），就能够观察到荧光管发光。磁场和电场都是向量场，它们是电磁场的两个方面。

一个描述电子和原子核所带电荷之间相互作用的基本理论，实际上就是一个描述电子和电磁场之间相互作用的理论。简单的量子理论为了解释原子结构，规定电子要遵守量子规律，但是却没有给电磁场加以量子要求，仍把电磁场看做是简单的经典场。尽管用这种方法能够在化学和原子物理学中得到许多很好的结论，但由于电磁场也应该遵守量子规律，所以它很明显不是一套完整的理论。毕竟普朗克是在电磁辐射现象中，首先注意到自然界的量子性质的。

那么电磁场在量子世界中是什么样子呢？电磁能是由许多很小的能量包（即量子——quantum）组成的，这些量子被称作光子，因此我们可以认为电磁场是由大量光子聚集而成的。因为我们没有根据假设所有光子的频率都是一样的，所以任何频率的光子都有可能出现。有时一些频率的权重会大于另一些频率，这就意味着具有不同频率光子的数目也不相同。这样，我们面前就展现出这样一幅图像：许多数量不等、具有不同频率的光子共同组成了电磁场。

我们可以用音乐作类比。我们对贝多芬及其他音乐大师推崇备至，但当交响乐队演奏他们的作品时，实际上只是在空气中形成了一些疏密相间的压力波——这种波被称作声音。一首乐曲是由许多不同乐器共同演奏出来的，每种乐器的数量都不相同，因此可以得到由不同频率的波组成的复杂声音。风格一致的乐章轻柔平和，在音乐厅内静静地回荡，中间穿插着富有节奏感的打击

乐器伴奏。与之相似，电磁场和电磁脉动也是由不同频率的电磁波共同叠加形成的。在这幅图景当中，光子就像是一只小号，或者其他某种音色嘹亮的乐器所发出的音色一样，掩盖了其他乐器的演奏。

电磁场和音乐之间还是有一些不同的，最大的差别就是一个是量子系统，而另一个不是：一个小号手可以把号声吹得更响一些，但一个光量子却不会变得更亮。在量子电磁场中，不同的频率表示具有不同能量的光子，某个频率的贡献比较大就意味着具有该能量的光子数目更多。可以和声波作一下对比，声波中某个频率的贡献更大仅仅意味着这个波的波幅更大，其中并不包含某种频率的贡献是由一些更基本的单位（量子）叠加而成的思想。

这样，光子的吸收和发射就更容易理解了。当一个光子被吸收时，就可以认为构成电磁场的光子数目又多了一个，就把电磁场中具有该能量的光子数加一；当发射一个光子时，构成电磁场的光子数目就少了一个；当跃迁的能级不同，所减少的光子的频率也不相同。这就是量子场论解释粒子诞生和湮灭的方式，至少在电磁场问题上可以这样解释。这也是狄拉克在 1927 年提出他的相对论性量子力学之前，引进量子场论的方式。

把场表示为许多符合量子规律的能量子的和，这叫做场量子化，偶尔也被不太确切地称作“二次量子化”。无论哪种说法，都表示正在进行量子化的是一个场，以区别于量子力学中对某个单个粒子（例如一个电子）运动的量子化。

转了一大圈，我们再回过头来考虑一下：既然场量子化可以很自然地解释光子的出现和消失，这种思想能不能应用到电子和正电子问题上呢？它们也会诞生或湮灭，就像电子偶的产生过程。另一个例子是放射性衰变，在该过程中电子和中微子会从“虚无”中诞生。因此量子场论不仅要解释这些现象，还应该能在同样的理论基础上推广到自然界所有的成员——像电子、正电

子、光子等。

量子场论应用到光子上的基础是电磁场的麦克斯韦方程组，正因如此，理论会自动地满足狭义相对论条件。用来替代“点状的”电子和正电子的场，也需要满足同样的条件。由于点粒子所遵循的狄拉克方程满足狭义相对论的要求，在此基础上建立的量子场论要引进一个新的场——狄拉克场，它满足场的狄拉克方程。就像电磁场是由光子组成的一样，狄拉克场也是其场量子——电子与正电子的集合。电磁场满足麦克斯韦方程组，狄拉克场满足狄拉克方程。例如一个电子和一个正电子对撞湮灭，产生一对伽马射线光子。这表示一个电子和一个正电子在狄拉克场的集合中消失，而在电磁场的集合中出现了两个光子。

这些量子场与经典的场差别很大。一个经典场——像风和温度的分布场以及用经典思想处理的电磁场，是平滑而连续地分布在空间中的。像“场在这一点的强度是多少”这样的问题有很明显的意义，但量子场却不是这样。尽管量子场也是分布在空间中的，但它既不平滑也不连续，当你靠得很近“观察”它时就会发现粒状的结构。观测的区域越小，场的强度就会起伏越大。这是由不确定性原理决定的：测量所间隔的时间越短，能量的测量值就越不确定；测量的位置越精确，动量的值就会越不确定。最多只能够算出场强度在有限时空区域内的平均值。

一个量子场是许多场量子的集合：每个粒子都是场中的一个起伏，即一个场强很大的区域。如果量子的能量位于不确定性原理所规定的狭小区域内，产生该量子的可能性会很高。粗略地讲，粒子大致相当于某个地方的旋风，它在该区域基本静止，但是这股旋风的能量是量子化的。添加了这个限制之后，我们就可以用这种图像来描述电子偶产生的现象了。一股旋风（即光子）与海洋的表面接触，在水面上形成一对小漩涡（电子和正电子）。一种场的量子湮灭了，另一种场的两个量子诞生了，这一切都是

通过两个场之间交换能量和动量实现的。

场的能量会随着场量子的增加或减少而变化。各种具有可能能量的量子对场的最终贡献，取决于它们诞生或湮灭可能性的大小。也就是说，如果产生具有特定能量的光子的可能性很大，那么具有这种能量的光子在量子电磁场中的数量也会很多。

处理量子化的一种方式，是将表示各种量子相对贡献的概率因子，都加以特定的量子规定。这种基于规则的方法称作正则量子化（canonical quantization）；这里，“正则”是指一般原理、法则或定律。按照 QED 理论中的处理方式，要先写出理论的经典主方程，然后用最小作用量原理改写狄拉克方程和麦克斯韦方程组，使它们能应用于各自相应的场。两种场都可以表示成所有可能量子的和；量子化步骤就是对概率因子要遵守的法则进行强制规定。

可是，当用这种基于法则的办法去计算物理上有用的结果时，却遇到了许多技术困难。值得庆幸的是，我们还有另一种方法——路径积分法可以使用。费曼在把量子力学应用于粒子物理时，首先引入这种方法，这是一次革命性的进展。在量子力学中，路径积分法可以给出粒子从一个量子态跃迁到另一个量子态的几率。在场论中，该方法变成了通过各种可能的干涉路径，计算场量子从一种可能组态跃迁到另一种可能组态的几率。

因此，量子场论建立在两大支柱之上，这两大支柱就是狭义相对论和宇宙的量子本质。最基本的实体就是量子的集合——场。粒子是场强起伏强烈的地方，对应于量子的定域结构。每种类型的基本粒子都对应着不同的场，例如光子场和电子-正电子场。运用该理论可以像处理光子一样处理其他玻色子，也可以用与处理电子和正电子相同的方法来处理其他自旋为 1/2 的费米子，因此量子场论比 QED 理论的应用范围更广。量子场论需要我们相信，我们周围坚硬的、存在相互作用的物质世界，实际上

只是少数的几种物质场，只是这些沸腾的场看起来感觉很硬。这是一幅令人惊异的图景，但似乎又是真实的。

量子场论不仅仅是一套理论，它包含整整一大类理论。一旦把整个骨架搭建起来，我们就可以布置一个个小房间了。留给我们的自由只剩下这些基本场所包含的内部对称性。对于参与基本作用的量子理论来说，规矩都是一样的：拿来量子场论，加上对称性，然后把它们搅和到一起。

粒子聚散

如果想知道量子场论为什么既要包含电子又要包含正电子，或者为什么要包含其他粒子及其反粒子，考察一下电子偶湮灭的过程，可以很好地帮助我们理解这一点。如果电子能够一个一个或成对地诞生和湮灭，就会产生电荷增加和消失的现象。而在自然界中没有发生这种事情，人们一直认为电荷是守恒量。这就意味着系统的总电荷应该是一个常数。因此，如果在某处有一个电子发生了湮灭，为了保证电荷守恒，肯定会有一个正电子同时发生湮灭。与之类似，电子与正电子也只能一同产生。实际上，一个有效的量子场论反映了狄拉克场方程的两个要求：电荷守恒，并且带电的费米子存在与其对应的反粒子。

另外，狄拉克在他最初的相对论量子力学的激励下，尝试用反粒子的思想来解释负能量。在场论中，不再出现负能量，但反粒子保留了下来，这是一个更加令人满意的结果。

但是，这反过来又说明，费米子具有一种奇怪而神秘的性质：不会有两个电子，或者两个其他任何费米子占据同一个态。这就像在一辆亚原子的公共汽车上，每个电子都会占据一个座位，因此其他同它一样的电子就不能占据这个空间了。也就是说不允许一个电子为了坐到比较靠前的座位，而坐到另一个电子的腿上，所有电子都必须坐到座位上去。尽管原子中的电子都喜欢

处于能量最低的态，但这条规则确保了每个电子都要处于不同的态。

从数学上来讲，由两个费米子 A 和 B 组成的系统，与把两个粒子对调后所组成的系统相反——即两个费米子形成了一个反对称态。换句话说，A 粒子结合 B 粒子等于负的 B 粒子结合 A 粒子。两个费米子交换满足反对称性。

我们举一个小例子，可能有助于理解这些问题。假设你正在开车去度假。旅途很长，车后座上坐着一对双胞胎孩子，他们分别坐在车的两边，你喂他们吃东西的时候他们总是哭闹。这时你把他们的位置调换了一下，神奇的是他们现在吃得开心多了。对一个偶尔路过的观察者来讲，孩子们的位置并没有发生改变——因为他们是双胞胎，外人无法分辨——但孩子们的最终状态（十分开心）却和初始状态（情绪不好）完全相反。

电子和其他费米子的这种反对称性质，意味着在同一处、结合在一起的两个完全一样的费米子与它们自己完全相反——A 粒子和 A 粒子结合等于负的 A 粒子结合 A 粒子，这是根本不可能的。这是一类新的场，与我们平常所接触到的标量场和向量场完全不同。

这种表面上看起来十分奇特的行为让我们感到，要想保留反粒子而去掉负能量似乎要付出很大的代价。实际上，抛弃负能量出现了这种问题之后，奥地利物理学家沃尔夫冈·泡利（Wolfgang Pauli）提出了著名而有效的假设，叫做不相容原理（exclusion principle）——在一个原子中，不会有两个电子占据相同的量子态。泡利在 1925 年初，用类似的说法提出了这个原理，那时现代量子理论还没有发展起来，电子自旋的思想也没有被提出。他最初的动机很简单：肯定会存在一些限制，来阻止原子中所有电子都挤在能量最低的那个态上。泡利知道，即使没有激发的原子，其电子也有一系列不同的能级，他的不相容原理是可以

解释电子为什么分布在各个能级的一种方法。因此，泡利不相容原理禁止电子（以及其他费米子）互相挤占各自的空间。

在那个时候，泡利对他的不相容原理只能给出这种解释：不相容原理只是一个使原子模型更加合理、与实验吻合得更好的专门工具。一段时间之后，泡利和其他物理学家才认识到，不相容原理可以通过在量子场论中保持能量恒为正而得出，这正是理论的量子本性决定的。

不相容原理，除了对理解原子中的电子以及整个元素周期表的本质具有重大意义之外，在理解最小层次上的物质结构方面也至关重要。不相容原理意味着，自然界中的粒子划分为两大阵营——费米子和玻色子。对于光子来说，不存在不相容原理——与电子相反，许多光子可以没有限制地聚集在一起。在亚原子公共汽车上的光子，可以一起坐在同一个比较靠前的座位上。在此基础上，我们可以很明显地看出，光子和电子的行为完全不同。例如在激光中，大量光子的行为方式完全一致；对于电子来说，即使不计电荷引起的效应，这也是不可能的。随着自然界中的基本粒子分为两大阵营，它们也分别满足两套用来描述粒子集团的“统计”系统：费米子满足费米-狄拉克统计，玻色子满足玻色-爱因斯坦统计。

费米-狄拉克统计是我们理解金属导电的基础，导电实际上是通过漂浮在金属中的自由电子形成的“电子气体”来实现的。从玻色-爱因斯坦统计可以推导出普朗克定律，从而解释与一个几乎完全封闭的盛有光子的盒子发光的相关问题（即黑体辐射现象）；玻色-爱因斯坦统计还是帮助我们理解激光现象的主要工具。

激光之所以如此特殊，是因为组成激光的光子都放弃了各自本身的特征，共同组成了一个单一的态，这只有玻色子才能够实现。该理论的一个结论就是：当激光束照到一个物体上时，所形

成的光点比普通光源所发出的光形成的光点要小得多。正是这种很强的聚焦性，使得激光有广泛的用途。

除了激光之外，玻色子的特殊性质还体现在其他许多方面。氦元素中丰度最高的原子是^{4}He，它也是玻色子。当把^{4}He冷却到大约－271℃，它就会变成超流体，也就是说它在流动的时候没有任何阻力。这是因为原子结合（或称浓缩）为同一个量子态。这种态的一个性质就是：具有超流动性的^{4}He能魔术般地爬上盛放它的杯壁。而非常罕见的^{3}He原子是费米子，它的行为与^{4}He截然不同。

1995年，埃里克·科内尔（Eric Cornell）、卡尔·韦曼（Carl Wieman）和他们在科罗拉多州玻尔德联合天体物理实验室的同事们一道，把铷原子气体冷却到绝对零度以上十亿分之几度——绝对零度是理论上能得到的最低温度，其值大约是－273.15℃。他们用一套磁场系统把原子束缚成一团云状，观察到在云状物的中心，原子的密度会突然显著地增大。他们所观测到的正是真正的玻色-爱因斯坦凝聚现象，而^{4}He因为有其他因素的影响，玻色-爱因斯坦凝聚现象没有那么纯正。铷原子形成了一团“原子物质”（atom-matter），这与平时常见的、由许多单个原子所组成的集团有很大差异。这是一种新的物质形态，在这种形式中，单个原子的特征都丧失了，就像在激光中光子放弃了各自的特征一样。同激光一样，这也是玻色-爱因斯坦统计的一个例子。科内尔和韦曼因他们的突出成果，与麻省理工学院的沃尔夫冈·凯特纳（Wolfgang Ketterle）一道，获得了2001年的诺贝尔物理学奖，凯特纳同一年用钠原子得到了类似的结果。

抛开量子场论，单从经验出发，将粒子按照自旋划分为两大类，其意义比从粒子计数统计推导出来更为深远。费米子是物质粒子，是构成宇宙的“原料”。玻色子是传递力的粒子。例如对于一个氢原子，其核心是一个质子（一种费米子），质子周围束

缚着一个电子（也是一种费米子），质子与电子之间的相互作用通过光子（一种玻色子）来传递。这是一个相当普遍的规律：费米子之间的相互作用是由玻色子传递的。

感受力

一个带电粒子是怎样感受到位于附近的另一个带电粒子的？一颗行星怎样对太阳的引力场做出反应？一个具有很强的相互作用的粒子怎样察觉到另一个粒子的存在？

在牛顿的世界中，力是一瞬间的事情。一颗行星轨道的变动会立即影响到另一颗行星。但对于爱因斯坦的理论，情况就不是这样了。一颗行星或者一个电荷的晃动，会引发另一颗行星或电荷的反应，但这种反应需要等到晃动以光速传播到另一颗行星或另一个电荷后才会实现。

这对于守恒定律来说有什么意义呢？能量或动量在两颗行星或两个电荷之间传递期间，会存在于两行星或两电荷之间的场中，这说明了为什么用场来处理问题是令人信服的。在量子力学中，这些携带力的场是由携带力的量子组成的。两个物体之间的相互作用过程，就变成了在两个物体间交换携带相互作用力粒子的过程。这是“交换力”（exchange force）的基本思想。

如果就像量子场论中所描述的那样，自然界真的允许粒子诞生和湮灭，这是否就意味着为各种粒子的消失和出现敞开了方便之门呢？简单地说，是这样的。除此之外，它还形成了交换力的基础。

由于电子带电荷，因此能够产生电场。用量子场论的说法，就是电子是一个光子源，电子可以通过产生和消灭粒子来发射和吸收光子，从而在其周围形成一团翻滚的光子云。两个电子可以通过交换与两个电子相联系的两团光子云中的光子发生相互作

用。这种交换光子的思想，取代了那种电场弥漫于两电荷之间的空间内的经典观念。

每个带电粒子通过交换光子也改变了自身的动量。这就表明它们之间存在一种微小的作用力。如果两个带电粒子交换大量光子，所有光子的共同作用效果就是在它们之间形成可观测到的吸引力或排斥力。

但是在交换光子这件事情上，还存在一些奇怪的问题。每发射或吸收一个光子都应该导致电子的能量和三维动量发生改变。但是对于那些普通的“真实”光子——例如像那些白炽灯泡发出的具有零静止质量的光子，如果把它们的动量和能量值加起来进行计算，就会发现交换光子的过程并不能满足能量和动量守恒。

最终的解释还是要依靠不确定性原理。当一个光子离开或到达带电粒子时，由于有短暂的时间间隔，它满足的能量守恒条件可以放宽。根据不确定性原理，能量守恒所能够放宽的限度，要依赖于在该时间间隔内的能量不确定度。借走的能量越多，借走能量的时间就越短。

与之类似，动量守恒也可以放宽，只要动量-距离不确定性原理允许。所借的动量值越大，借走动量的过程所经历的距离就越短。

如果探测到满足放宽了的能量、动量守恒条件的真实量子，就意味着违反了神圣的能量和三维动量守恒定律。认识到这种困难，就会明白参与这种“粒子交换”的粒子，不会是那种普通意义上的可观测粒子。物理学家们把它们称作“虚粒子”（virtue particle）。虚粒子是一种彻头彻尾的量子现象。

因此，量子场论为相互作用力提供了一整套新的诠释，即力是交换携带相互作用的粒子的过程。带电粒子之间通过交换光子实现相互作用。引力场所对应的量子是引力子（graviton），物体之间通过交换引力子产生互相牵拉的万有引力。弱相互作用是通

过交换 W 粒子和 Z 粒子实现的。

携带力的候选粒子

用交换光子流来解释带电粒子之间相互作用的思想，是由费米（Fermi）和德裔美国物理学家汉斯·贝特（Hans Bethe）在1932年提出来的。3年之后，日本物理学家汤川秀树（Hideki Yukawa）将这种思想发展成第一套有关强力的场论。

汤川秀树根据强力只有在原子核内部才起作用这一点，推断出它的作用范围很短。因此，携带强力的交换粒子应当具有比较大的质量。根据不确定性原理的规定，粒子的质量大，就意味着在交换粒子时只能在很短的时间间隔内借走能量，在这么短的时间内粒子只能运动很短的距离。因此，汤川秀树假设的那种新交换粒子的质量值，应该和强力的作用范围——大约是质子半径——直接相关。对于2飞米长的距离（1飞米是 10^{-15} 米），汤川秀树交换粒子的质量应该是大约100兆电子伏。因此不足为奇，当两年后人们发现一个质量为106兆电子伏的粒子时，立刻断定它就是汤川秀树所预言的携带强力的粒子，这也将汤川秀树的思想带入了人们的视线当中。

汤川秀树十分幸运。现在已众所周知，卡尔·安德森（Carl Anderson）于1937年所发现的粒子实际上是 μ 子(muon)，是以前所介绍过的电子的表兄弟，它实际上与强相互作用无关。但不久，汤川秀树预言的携带强力的粒子也被发现了，被命名为 π 子(pion，或称 π 介子)。不带电的 π 介子质量大约为135兆电子伏。带正负电荷的 π 子质量更重一些，大约是140兆电子伏。

π 子交换也被用来描述强力在稍微大于1飞米的距离上是怎样实现的。这就是说，交换 π 子更应该是原子核物理的内容。在原子核物理中，理解质子和中子怎样结合为原子核是一个艰巨的任务。物理学家们还认识到，π 子本身实际上也不是最基本的粒

子，它也是由更小的组成部分构成的。这些更小的组成部分怎样形成π子？这是由QCD理论描述的。

π子并不是携带力的唯一的大质量粒子——弱相互作用的力程也非常短，因此弱相互作用也是通过大质量交换粒子——W粒子和Z粒子来传递的。QCD理论是关于强力的理论，由于强力是短程的，QCD理论可能也需要包含较重的携带力的基本粒子。事实是这样吗？答案是否定的！QCD理论中确实存在携带力的基本粒子，但它们却是没有质量的。QCD理论在解释强相互作用的短程性时，有一些新特点。

虚粒子

在经典理论或是相对论量子理论中，电子都是一个单独存在的孤零零的物体。但量子场论却认为：实际上在电子的周围有大量的粒子离开电子或被电子吸收。支配这团“虚云雾”的是光子，这些光子靠借来的能量和时间存活，并且产生了经典理论中叫电场的东西。同样，光子周围也存在由一些匆匆而过的虚粒子——电子和正电子所组成的云团，电子和正电子反过来又成了其他虚光子的发源地。

当电子撞击质子，两个粒子之间通过交换虚光子产生电磁相互作用。看起来就像：围绕在电子周围的虚光子都在四处搜寻，最终它们在带电的质子或其他带电物体所呆的地方找到了它们的新家。因此，这个碰撞过程是由极短时间内并不真实存在的虚粒子传递的，这些虚粒子只遵守符合不确定性关系的能量和动量守恒定律。

当实验者们将电子和质子碰撞，他们观察到的只是电子被散射了——它的方向发生改变，或许能量和四维动量都有所减少；质子也发生了一些反冲。电子初末四维动量的改变，经由虚光子传递给了质子或质子的组成部分。由这些光子所携带的四维动量

称作动量传递。

通常，一个粒子四维动量的平方就是与该粒子相关的静止质量，在这种情形下它自动满足正值原则。但在电子和质子碰撞过程中，动量转移的平方却得到了负值。因此，高能物理学家们将动量转移平方取反，最后得到了一个正的量。他们把这个量记作Q^2，它是一个在粒子物理中，每次讨论粒子碰撞时都会用到的标准量。

一个从白炽灯泡发出、在屋子里乱窜的真实光子，其四维动量的平方是零，因为它的静止质量就是零。另一方面，一个在瞬间借到能量的虚光子变得具有质量了，所以它的四维动量的平方不是零。其值是负的，这对应着它的非物理本性，没有人能“看到”虚光子。

因此，出现在碰撞过程中的虚光子会具有各种各样的质量，但是只能在很短的时间内存在。虚粒子的质量越大，它的寿命以及它所能影响到的范围就越短。正因为如此，动量转移越大，光子就会越深入到目标的结构当中去。在效果上，动量转移就等同于电子散射“显微镜”的分辨能力，动量转移越大，所能探测到的细节就越细微。

这就是为什么Q^2（有时就简称为动量转移）在粒子碰撞的实验数据图和理论物理学家们的计算中频繁出现的原因——它是粒子加速显微镜放大能力的度量。随着放大倍数的增加，实验者们就能更清楚地观察物质的结构。

得到一些结论

当计算太过复杂，理论物理学家们不知道怎样算出结果的时候，怎么办？可怕的是，这是他们经常遇到的问题，而不仅仅是偶尔出现的特殊情况。一个真实的物理问题，通过处理后最终能

得到一个简洁、确切的数学方程（数学家们称之为封闭形式的解析解），这种情况是很罕见的。大多数问题都太难处理了。

那么当问题太复杂的时候该怎么办呢？除了全部放弃和更努力地直接解决问题之外，还有三个基本方法。第一个方法是解决一个比较简单的问题，例如把原来的问题先简化一下再处理。第二个方法是用电脑处理原来的问题，这称作模拟方法。第三种是近似方法。

解决一个比原始问题简单的问题有什么意义呢？其意义就在于，在简单的情况下很可能会找到解决真实问题的思路，有时甚至能从简单情况直接推广到真实的物理问题。或许能从中看出一些物理方面的性质，这是我们解决问题的最终目的；也有可能会发现解决真实问题的方法，或者至少说明了这个问题为什么不能解决。

这不仅要求能够解决较简单的问题，还要求掌握找出较简单问题的技巧。从真实问题中抽象出来，同时保证物理可信度的较简单问题称作模型。物理学中到处都是模型和简化过的问题，其中有许多十分著名。例如可以通过假设原子核的质量是无穷大来简化问题，之所以能这样做是因为原子核的质量与电子相比要大得多。在研究大量物质时的一个重要模型是著名的伊辛模型（Ising model）——伊辛模型把实际问题简化为一个空间点阵，每个格点上的自旋之间存在相互作用；理论物理学家们通过伊辛模型可以推导出严格解。

现如今，电脑计算能力强大而廉价，这为我们拓展了一片新天地。一堆沙子之间的相互作用力十分复杂，其复杂程度很难想象。尝试理解这个过程的一个方法是，将有关沙粒滑动以及互相挤压的所有基本力学规律都输入电脑中，然后让电脑一步一步地来处理成千上万的方程，从而建立一个真实沙堆的模型。这个问题是计算机模拟的一个例子，是那些对任何东西（如谷粒、铁矿

石或其他东西）形成的堆状物都感兴趣的人经常做的工作。电脑强大的计算能力也为 QCD 模拟（称作格点量子色动力学，lattice QCD）展开了广阔的现实前景，这是第十章的内容。

然后，还有近似的方法。如果问题特别的难，可以利用一些数学技巧，给出足够精确的近似解来。在 QCD 理论中比较重要的一种数学技巧就是微扰论（perturbation theory）。微扰论不是一个新的理论，也不是 QCD 理论的固有部分，甚至都不是粒子物理所必须包含的部分。实际上，它所产生的年代和数学本身一样久远，它的应用范围也同样广泛。它起源于对行星运动的研究，是由像艾萨克·牛顿、约瑟夫·拉格朗日（Joseph Lagrange）、威廉姆·哈密顿（William Hamilton）等其他一些为基础数学的发展和完善作出突出贡献的人发展起来的。

行星运动与其他问题一样，是一个全面了解微扰论的不错的例子。例如一颗绕地球旋转的卫星的运动，由于地球在绕赤道的方向上有些“发胖”，卫星的轨道不同于简单的圆周轨道。在这个卫星绕地球旋转的例子中，微扰论需要把描述的过程分为两部分。一部分包括绝大部分相互作用，在此条件下，卫星绕一个完全为球体的地球运动，该问题可以通过某些数学方法解决。另一个部分要小得多，代表了简单情况与现实情况的差别，即微扰。数学家们无法用简单而直接的方法解决这种问题。在卫星的例子中，微扰是由于地球不是完美的球体在卫星轨道上引起的效应。

一个经典例子就是月球绕地球运动，而地球又绕着太阳运动。写下并解出描述这三个天体运动细节的方程，听起来并不是一个十分离谱的要求，并且对预言潮汐很有帮助。

这个问题是一个典型的三体问题，这意味着三个物体中的每一个都与其他两个有相互作用。这类问题一直困扰着我们，直到 1994 年，美国数学家夏志宏（Zhihong Xia）才证明了三体问题不能直接解出。实际上，夏志宏证明了三体问题是无秩序的，这

就意味着即使给定了地球、太阳和月亮现在的确切位置，也不可能精确地预言它们未来的准确位置。

这就是数学的限制，尽管在这个例子中，强大的数学至少能解释这种困难。现在，数学家们也理解了，为什么微扰方法在以前的天体力学中是如此的重要——根本不可能有严格解，而电脑在那时还没被发明出来。具有讽刺意味的是，夏志宏正是用微扰方法证明这一点的。

其实微扰论的思想，就是把问题分成两个部分。其中一部分是将系统加以简化，使其满足现有的数学工具。另一部分表示简化后的情形与原问题之间的小偏差，这个微小的偏差会使问题更接近现实，但也更加难以解决。这种划分是处理问题的第一种技巧。

第二个技巧是要得到一个实际答案。数学家们通常用他们称作“多项式展开”（series expansion）的方法来解决这个问题。多项式展开就是将一个数学表达式表示成一系列指数项相加的形式。例如，表达式 $(1+x)^{20}$ 可以表示成一系列指数项的和：$1+20x+190x^2+\cdots$（其中……表示后面有更多的项）。对于某些 x 的值，前三项给出的值与真实值十分接近，但有些情况下却不是这样。例如，如果 x 是 0.01，前三项相加的值是 1.219，而真实值比 1.220 稍微大一点点。如果 x 的取值更小，结果就会符合得更好：如果 x 取 0.001，级数的前三项的和与真实值吻合到小数点后 5 位。但是如果 x 取更大的数，情况就变得糟糕了。当 x 取 0.1，真实值大约是 7，而级数前三项的和却不到 5。要想让结果吻合得更好，最简单的方法是多写出一些级数的项。

这就给出了一个重要的推论，也是这种方法的一个主要规则：写出的项数越多，结果就吻合得越好，这起码在以上例子中以及绝大多数情况下是成立的。上面所用到的级数的例子带有人为的性质，而并不是必须这样处理，因为该级数是有限的，即我

们可以在写出它的所有项后停下来。最终的结果包含 21 项，而这正是我们把（$1+x$）自乘 20 次后，进行代数展开所得到的结果。如果把这 21 项全部用上，无论 x 取什么值，我们都会得到精确的结果。

然而，作为数学上的通用表达形式，级数一般都是包含无穷多项的。数学家们有一套用来得到所需级数的规则，级数的前几项通常会给出与真实结果极为接近的值。如果你还需要更加精确的值怎么办？没问题！你只需要再向后多算几项就可以了。

微扰论用到的最典型的数学工具就是级数展开。例如，在卫星围绕变形的地球运动的问题中，最终结果就是用级数展开给出的。第一项就是卫星绕球体运动的简单情形。第二项以及后面的项，则对由地球形状的不规则所引起的轨道偏差逐级做了精确的修正。级数是用问题中某个参数幂次逐渐增加的乘方项表示的，该参数在微扰展开中称作展开参数。当物理学家们讨论更高次的修正时，是指他们正在计算含有展开参数幂次更高的微扰项。

级数中的每一项都应该比前一项小，粗略地说，其中隐含着对前一项的限制。如果不满足这个条件，后面的项就不会给出逐步精确的结果，算出的结果会非常离谱。由于级数序列没有尽头，因此就无法找出一个有意义的结果。因此，展开参数一定要很小，一个小的量的次数越高，所得的值就会更小。

如果后面的项减小得特别快，则只需要级数的很少几项就能得出十分精确的结果。由于后面的项可能算起来很麻烦，因此这是一件很好的事情。但是，有利必有弊。如果展开参数很大，比如说接近 1，那么次数更高的项就会减小得特别慢，会和前面的项差不多大。如果在微扰论中出现这种情况，或者至少在利用微扰论时出现这种情况，那就无法得到有用的结果。对于 QCD 理论，这是一个很实际的问题。

在 QED 理论、电弱理论和 QCD 理论中，微扰论是除实验之

外获得结果的主要工具。方法就是将主方程划分为两个部分。一部分表示理论中自由运动的粒子，可以直接地解出来。另一部分是微扰部分，表示粒子之间的相互作用，这部分很有意思。这一部分也正是结果中存在微扰项的原因。

对于 QED 理论来说，相互作用的部分比较简单，就是光子和电子-正电子之间的联系。其中所有的相互作用因素都是由电子电荷引起的。单凭直觉也可以感觉到，电子所携带的电荷表征了电子受电属性影响的大小，它必定控制着相互作用力的强弱。而在微扰方法中产生的一系列越来越小的项，对应着相互作用的精确度越来越高。正如我们马上就会看到的，这告诫我们不可能获得最终的精确结果。

由于电子电荷 e 只在 2 次、4 次、6 次以及更高的偶次项中出现，作为惯例，物理学家们引入了一个新的物理量——耦合常数（coupling constant），它的值与电子电荷的平方成正比。在最终的结果中只会出现耦合常数（用字母 α 表示）。振幅的微扰级数后面的幂次对应着一个有关耦合常数次数更高的级数。在粒子物理学中用到微扰理论时，耦合常数是作为展开参数出现的。在 QED 理论中，微扰级数表示为 α 的一次幂乘上其他因子（一次项），后面是 α 的二次幂乘上其他因子（二次项），等等。

从 QED 理论、QCD 理论或电弱相互作用理论的主方程出发，找出一个可行的微扰论解决方案是一项庞大的工程。幸运的是，这项工作在很久以前已经完成了，在大多数粒子物理课本后面，附录中都用规则的形式列出了结果。当理论物理学家们想要进行计算的时候，他们不必每次都从主方程出发，直接从这些规则开始就可以了。这些定则称作费曼规则，是用费曼图的形式给出的。

一位名垂千古的面包师

实验物理学家们喜欢将理论物理学家们的预言同自己的实验结果作比较，从而寻找理论与实验的一致之处（或不同之处）。典型的实验是将一束粒子射向靶粒子，然后测量从碰撞点产生的粒子的数量、方向和性质。

实验物理学家们最常测量的典型物理量，也是理论物理学家们要相应计算的量，就是碰撞截面。碰撞截面具有面积的量纲，它表示了粒子之间相互作用的强度。

计算碰撞截面要有三个步骤。第一步要计及碰撞粒子的数量和速度——这是一种粒子流效应。第二步要规定能量和动量在碰撞碎片间分配的方式。这些都是例行的程序，真正的工作实际是第三步，也就是最后一步：所讨论的相互作用发生的可能性。

在量子场论中，粒子可以从一无所有的空间（物理学家们称之为“真空”）中产生，同样也可以消失在真空里。如果产生了一个粒子，但稍后又在空间中不同位置处消失了，那么粒子就穿越了两点间的间隔。描述该过程的量子振幅有一个名字：格林函数（Green's function）。由于它建立了粒子在时空内的某一点产生以及在另外一点消失之间的联系，实际上应该被称作是“两点间的格林函数”。它有效地描述了粒子从一个点传播到另一点的过程，因此格林函数也被称作传播子。在 QED 理论中，就存在描述电子或正电子在一个地方产生，最终又在另一个地方湮灭的传播子。同样也存在另一种描述光子的传播子。在电子偶湮灭现象中，一个电子和一个正电子相互碰撞发生湮灭，并在碰撞处产生两个光子。由于这个过程涉及了四个粒子，该过程的量子振幅是由四个点的格林函数描述的。

正是这个叫做格林函数的工具，描述了粒子散射的物理核心。要最终计算出碰撞截面，就要计算出格林函数。计算格林函

数的最好方法就是利用费曼图——费曼图是从费曼的路径积分理论中抽象出来的。

由于格林函数是量子场论的核心，你一定认为它是以一个著名的量子场论大师的名字命名的吧？但实际不是这样：格林函数是以乔治·格林（George Green）的名字命名的，他是英国诺丁汉（Nottingham）的一位磨坊主兼面包师，生于 1793 年，卒于 1841 年——比量子场论的诞生早了将近 100 年。格林是一位自学成才的数学家，他同其他数学家一道，提出了一套可以有效解决电磁学以及热流问题的数学工具，也就是现在我们所熟知的格林公式。

一个简单的格林公式，比方说一个没有相互作用的、从一个地方运动到另一个地方的传播子，是很容易写出来的。用来描述没有相互作用系统的四点格林函数，其实就是一些描述从起点运动到终点但感觉不到对方存在的粒子的两点格林函数。

一切看上去索然无味，只有加上相互作用才会有趣。正是相互作用导致了实验物理学家们可以观测到的碰撞截面，但将粒子联系在一起的相互作用也使得计算变得特别复杂。实际上，计算两点间或四点间的格林函数不再那么简单了，因为你不能把它们独立地移到方程的一边，它们总是要用它们自身来表示。这就说明了为什么严格解决量子场论问题是不可能的——没有人知道该怎样处理这些数学问题。人们所知的，计算感兴趣的格林函数的唯一方法，就是在微扰论的指导下运用近似手段。

在 QED 理论中，结果就是一系列（裸）电子-正电子和光子的传播子，因电子电荷连接到几个顶点上。例如，一个完整的、用来描述正在运动且具有相互作用的电子-正电子对的格林函数，就是一系列由耦合常数组成的幂次递增的项：第一项，也就是最简单的一项，对应着一个无相互作用的电子从一个点运动到另一个点。接下来的一项有两个交汇点（即顶点，因而耦合常数的次

数是一次的)，对应着一个电子从 A 点运动到 B 点，一个电子协同一个光子从 B 点运动到 C 点，以及一个单独的电子从 C 点运动到 D 点。第三项中，耦合常数的次数增加到二次，对应着更为复杂的解释。

这些微扰级数项可以用一个图表画出来，一切都会变得突然清晰起来。图 3.4 所示的就是一些 QED 理论中，费米子所有的两顶点格林函数的贡献。这些都是著名的费曼图的例子。

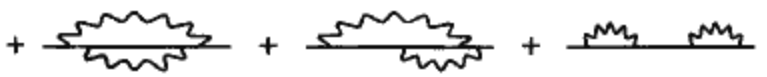


图 3.4 电子的两点间格林函数。完整的电子传播子——那条很粗的实线是无穷多微扰级数项的和，这些项在这里用费曼图表示

图 3.5 显示了另一费曼图——电子-正电子散射过程的费曼图（$e^+ + e^- \rightarrow e^+ + e^-$），因此它们表示了四点 QED 格林函数精确到耦合常数一次项的情形。图中只有两个费米子-光子顶点，所以在表达式中出现的耦合常数是一次的，这些四点方程的最简单项表示了电子和光子的相互作用。图 3.6 表示的是同一过程二次项的费曼图。

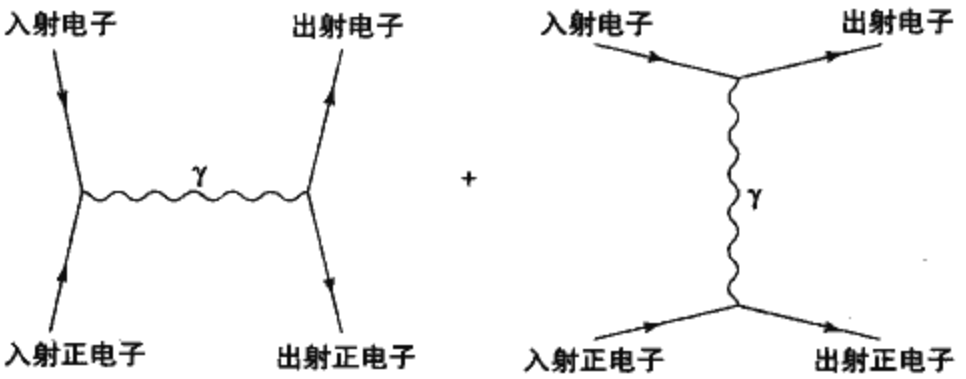


图 3.5 对电子-正电子散射有贡献的一级费曼图

费曼图

费曼图第一次出现在 1949 年，费曼在一次量子电动力学讨论会的论文中，用它来表示所讨论的理论中的微扰级数展开项。它们不仅仅是一些简单的、描述粒子碰撞的怪图，图中的每一条线、每一个点都对应着精确的数学表达式。按照惯例，直线表示费米子传播子，在 QED 理论中，它们对应着经过的电子或正电子。波浪线表示光子传播子，它们同表示费米子的直线有交汇点，称作顶点（vertex）。在每一个顶点处，四维动量都是守恒的，这表示流入顶点的四维动量和流出顶点的值应该相等。严格来讲，表示费米子的线应当有表示方向的箭头。没有圈的图被称作树状图，这是因为图中只有简单的分叉；含有圈的图在末端会有树状图所没有的小刺。

所有可能存在的、需要实际计算得出的图，都可以用简单的“子图”拼出来。在 QED 理论中，这种子图总共只有 3 种，分别是电子传播子、光子传播子和有 3 个分叉的电子-光子顶点。除了这些以外，还要加上一个基本规定：所有各种可能的费曼图都有一定数目的顶点。另外再加上一些技术细节，就足以完成一个对粒子散射过程的实际计算。

费曼的作图方法将微扰论计算大大简化了，几乎将这种工作变成了一种体力活。大学毕业生们即使没有听说过格林函数或路径积分，也可以用这种方法算出有意义的结果。现在，几乎所有人都能完成这项工作了！

当然，并不完全如此。实际上，对于包含大量顶点的费曼图来说，随着顶点数量的增加，计算的工作量会猛增。连同其他一些问题，决定了实际上只有最简单的情形才能够计算出来。实际上，人们从来没有完整计算出包含超过 4 个圈图的费曼图的情况。例如，将电子的反常磁矩计算到 3 个圈图要包含 72 个图，计

算到 4 个圈图则要包括 891 个图，如果到 5 个圈图，这个数量已经相当惊人了——要计算 12672 个图。我们可以展望一下前景：在经过 3 年的努力后，人们终于在 1996 年征服了含有 3 个圈的图，含有 4 个圈的图的计算过程还在进行中。在本章开始引用的理论值 a_{theory} 仅仅估计了 4 个圈的贡献。

靠费曼图来处理问题还存在理解上的困难。这看起来十分明显：例如图 3.5 中的某一张图，电子和正电子的散射过程，看上去就像是电子和正电子先结合成一个光子，然后再经过电子偶产生过程重新出现，就像光子变成了一个电子和一个正电子。尽管这种低次项的费曼图可能同真实的物理过程十分接近，但总的情况却不是这样。费曼图是级数展开，而简单的树状图只是展开的前几项。在现实中，物理过程实际上是由一团光子、电子以及正电子形成的云雾共同传递了相互作用。

即使是一个简单的电子从一点运动到另一点，也能说明这个问题。电子从一点传播到另一点的振幅包括一系列更加复杂的后继项，而并不仅仅是一条电子线（图 3.4）。从整体上看，它是一个“穿戴整齐”的传播子，而不是一个“裸”电子，这种含有圈的费曼图称作电子的自能图。在费曼图中，这表示了量子场论的物理思想：像电子这样的粒子是同一团光子云相联系的。“真实”存在的是穿着衣服的传播子，而不是光着身子的。

从一个地方运动到另一个地方的光子也是这样。一个穿着衣服的光子传播子包括“裸”传播子，然后是形成一个费米子圈的一次项，更高次项的图包含更多的圈。这种圈图被称作光子的自能图，或者叫真空极化图（因为光子在通过真空时会扰动真空）。这是一种纯量子效应，在经典物理中没有对应，但可以通过实验观测到。

难以琢磨的展开

QCD 理论中的微扰论同 QED 理论几乎完全一样。但有一个

主要的差别，它从某种意义上来说，成为了 QCD 理论中阿喀琉斯的脚后跟①。尽管在电动力学中，通过费曼图可以方便地使用微扰论，并且结果也符合得很好，但在 QCD 理论中却不能这样。QCD 理论存在一个问题。时常会有这样的情况：一些很小的修正项，最终会变得比它们要修正的东西还要大，就像在大海的波浪中泛起的涟漪，结果比波浪本身还大。从数学上讲，随着计算精度的提高，展开参数的次数也会增加，之所以产生这种问题，实际上是由于展开参数不够小。一旦发生这种问题，理论物理学家们就有了大麻烦。于是，有时要求助于其他数学工具，有时要求助于计算机，但是没有在理论上解决这个问题的严格方法。在一些情况下，不能运用微扰论成为了粒子物理学中一个重大的困难。

但这还不是最糟糕的，更加麻烦的是，微扰论本身也有本质上的问题。这种问题与前面提到的 QCD 理论的微扰论困难不同，在日常的计算中不会出现，但它却腐蚀了整个量子场论大厦的根基。将微扰论运用到场论中会得到逐项递减的级数，理论物理学家们一度假设这些级数拥有一种称作“收敛性”（convergence）的性质。但在 1951 年，英国理论物理学家弗里曼·戴森（Freeman Dyson）指出这并不是普遍成立的，这似乎是一个技术问题，但具有深远意义。

除了在微扰论中出现的那些级数之外，数学家们还了解许多其他的级数。他们用级数项收敛的说法来表示一个级数逐项减小，并且把所有项加起来最终可以得到一个数值——级数和。例如级数 1＋1/2＋1/4＋1/8＋1/16＋…（后面有更多的项）是收敛的，这个无穷多项的序列加起来得 2。

---

① 译者注：阿喀琉斯（Achilles）是希腊神话中的英雄，小时候被其母亲倒提着在冥河水中浸泡过，因此除了脚后跟之外全身刀枪不入。这里喻缺陷。

在场论中运用微扰论，最好能给出一种称作渐近级数的级数，这是一个很令人头痛的问题。为了说明其中的困难，我们可以举一个渐近级数的例子，它的前几项［该级数的一般项表示为$(-1)^{n-1}(x-1)!/x^n$，其中 $x=2$］是 $1/2-1/4+2/8-6/16+\cdots$，等等，由于后面的项越来越小，所以看起来很不错。但当我们写出更多的项时，问题就出现了：$+24/32-120/64+720/128-\cdots$这个级数所有项的和是无穷，当把所有项都加起来之后，级数不是收敛到一个数，而是发散的。

比较严格地讲，如果在展开系数趋近于零时，级数所表示的量与级数前几项之和的差也趋近于零，数学家们就称这种级数为渐近级数。对于展开参数大于零的情况，结果最后总会出错。也就是说到了某些点，级数某些项的大小会突然增加，级数就会发散。

在量子场论中，展开系数是相互作用的耦合参数，它很明显不是零，因为耦合参数是零就意味着没有相互作用。而耦合参数取有限值就说明在微扰级数的某些点上，级数某一项的值会比前一项更大，而不是更小。接下来的项会越来越大。

因此在现实理论中，微扰论并不是无限精确的。即使有最强大的计算机和最庞大的理论物理学家团队来计算大量的费曼图，也不可能计算出任意精确的结果。微扰论存在一个精度极限，该事实直接导致了微扰级数是渐近级数而不是收敛级数。对于一套像 QCD 这样的理论，即使是在理想条件下运用微扰论，我们也只能得到近似的结果。

QED 理论是迄今为止所知的最完美的科学理论，但从严格的数学观点来看，它仍毫无意义。QCD 理论和电弱相互作用理论一样，不可避免地有同样的问题。这就是生活！

事实上还存在其他问题。对于一些理论，尤其是 QED 理论和科学家们的“玩具”理论（尽管不是在 QCD 理论中），微扰论

在其他某些情况下也不成立。耦合常数随着能量的增强会慢慢变大，等达到足够大的能量，耦合常数就不再是小量，微扰论也不再成立。很幸运，对于 QED 理论来说，实验中所涉及的能量都足够小。理解这个问题的关键，同时也是 QCD 理论的一个关键性质，是 20 世纪数学物理领域的最重要思想之一：重正化群思想（renormalization group）。

驯服无穷大

费曼曾经把它描述为“欺骗”以及“脑筋不正常的过程”。有一些人甚至把它归为数学欺诈。它也曾经被吹捧为理论物理学家们建立的认识宇宙的指导准则。直到不久前，理论物理学家们才开始把它同某些问题联系起来，实验物理学家们早已了解这些问题。在最原始的层次上，它其实是一种从爆炸式增长的费曼图中找出有意义结果的方法。这就是重正化（renormalization）。

问题出现了。对在电子或光子周围乱窜的虚粒子来说，它们的能量没有限制。假设它们遵从海森伯的不确定性原理，它们所借到的能量与它们借的期限有关，那么虚粒子就可以借到它想借到的任意能量。这就会出现超出想象的庞然大物，这些虚粒子可能比质子这样的粒子还要重。这样重的粒子只能存在极短的时间，它所能涉及的长度范围，即使同原子核尺度比较，也是非常小的。因此，量子场论要求在计算中包括在无穷小尺度内存在的、具有无穷大能量的虚粒子。理论物理学家们这样做时，会得出无穷大的结果。

在费曼图中很容易看出问题所在。在几乎所有量子场论的基本计算当中，理论物理学家们都会遇到包含封闭圈的费曼图。例如电子-正电子散射过程：$e^+ + e^- \rightarrow e^+ + e^-$，最简单的计算只包括树状图（图 3.5），但是接下来的图中就包括圈了（图 3.6）。

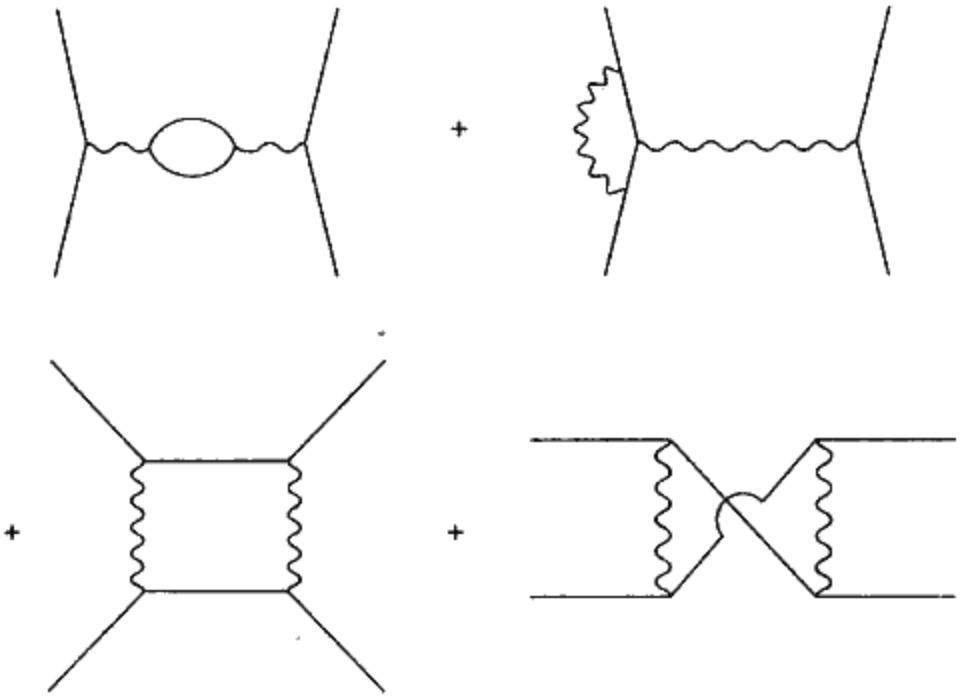


图 3.6 对电子-正电子散射有贡献的单圈费曼图（或者称二级费曼图）

在费曼图的每个顶点处，四维动量的总和应该守恒——进入顶点的四维动量值一定要等于离开顶点的四维动量值。当顶点扩展为一个圈，对于四维动量会出现一些奇怪的现象。要使四维动量的总和守恒，看起来唯一的方式就是只能在闭合圈中引入四维动量，而在其他任何地方都不行。反过来讲，这又表示可以假设圈内的动量为任何值，只要在顶点处符合要求即可。图 3.7 显示的是一条存在一个圈的光子线的费曼图，该图能很好地说明这个问题。

计算各个费曼图贡献的法则为，圈中所有动量的贡献都要计算到总结果中来，这就是症结所在——圈中的动量总是在源源不断地流淌着，这些闭合的圈导致了无穷大的结果。计算所有贡献的结果，就是得到一个没有任何意义的无穷大。只有有限的结

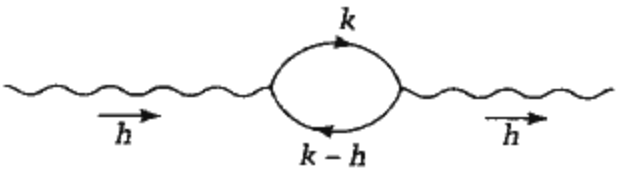


图 3.7　光子动量 h 经过图中的圈，没有任何限制电子动量的机制

果，也就是有实际意义的最大数值结果，才能用来与实验相比较。我们马上就会看到，理论物理学家们认为，量子场论仅仅是某个更基本的理论的近似，而无穷的出现只是这种近似的反应，这种理论上的 180°大转弯为现代的观点奠定了基础。因此，人们不再把无穷的出现当作是不正常的现象了。但这是 20 世纪 70 年代以后的事情，在此之前，无穷大一直让全世界的理论物理学家们头痛不已。

这种趋于无穷的过程称作发散（divergence）。实际上，这是一个紫外发散的例子。“紫外”是指发散的“源头”存在非常大的动量，或者对应着非常短的波长——换句话说，潜在的大质量的虚粒子“儿女”们只在无穷小的长度范围内起作用。把它命名为紫外发散可以方便地将其与其他发散区别开；那些在低能环境下出现的发散稍微容易处理一些，它们称作红外发散。

怎样处理带有圈的费曼图中产生的无穷大问题已经困扰了 QED 理论好多年，这几乎迫使人们放弃量子场论。同样的发散问题也出现在 QCD 理论与电弱相互作用理论当中，对所有理论来说，处理该问题的方法都是一样的。

这种修正方法叫做重正化。重正化将无穷问题隐藏起来，并提供了一套能得出与实验比较的、有意义结果的方法。重正化的物理图像，与一个在固体中运动的电子类似。假定有一个电子，可以测量它的质量，并且确定测量结果与课本中给出的数值相一致。同一个电子如果在固体中，会与由正离子组成的晶格和其他

电子发生相互作用，因此它的行为使它的质量看起来与书本上给出的不同。同样，固体中电子的有效电荷也被环境修改，与书上给出的值不同。当电子陷入固体当中，它所在的环境就会对它的质量和电荷进行修正（或称“重正化”），使这些值与书中给出的值不同。

在量子场论中，这种思想发展为电子与它所在的真空发生相互作用，这就意味着真空并不是一个没有任何作用的空白空间。因此，电子周围就包上了一层由虚光子和其他粒子组成的云雾，所以在现实中不可能看到“赤裸”的电子。一个测量电子质量和电荷的实验，实际得到的是由于粒子云的存在而被修改过的值。

这同在固体中运动的电子有两个关键差别。第一，固体中的正离子和其他电子只能有限地修改电子的质量和电荷，但在真空中裸露的电荷和质量的修改必定是无穷大的。第二，尽管一个电子可以在固体中移动，并且能孤立地检验，但电子永远不能从真空中孤立出来——由真空引起的对裸露量的修改永远无法消除。

重正化过程分两步。第一步，一定要用一种可以比较、加减以及进行其他操作的形式将发散写出来，就像有限的情形一样。由于数学家们普遍对无穷大加无穷大，或无穷大减无穷大这类问题的意义表示怀疑，因此这方面还需要做很多工作。第二步，要把各种各样的发散一起考虑，从计算的其余部分抽出来，最终留下一个物理上有意义的结果，用来同实验比较。这种方法的第一部分称作规范化（regularization），第二部分是重正化。可以采用无穷多种方法来完成这件事，但最终得出的所有结果都必须相同。毕竟一切都要与现实世界相符。

修剪时空

那么先看第一步：规范化。问题的根源在于可能产生具有巨大能量的粒子，或者说是理论在小尺度上的性质。用费曼图来说

明，问题就在于圈中的动量能够增加到无穷大。解决这个问题最简单的办法就是直接处理最明显的问题，在计算中对允许的最短距离加以限制，或者等价地在计算中规定一个虚粒子质量的上限。这就是费曼在20世纪40年代后期，在处理电动力学发散问题时所用的方法。他在方程中引入了一个符号代表可能出现的最大动量值。实际上，他没有设定任何确定的值；这样做就表示动量被“截断”了，其效果是能够用该符号表示出答案来。如果将真实值代进去，答案就又变成无限的了。

有时，这种方法效果相当好。当基本问题本质上是截断的时它就更有意义，例如在固体理论中，原子之间形成的自然晶格空间就造成了动量截断。像QCD这样的理论所描述的平滑的时空，实际上就是这些晶格间距趋于零的结果，也就相当于截断项趋于无穷。在实际操作中这种方法是毫无用处的，这对于QCD理论也毫不例外，因为它不能满足理论的对称性。

在QCD理论中存在一套规范化方法，可以为其他各种方法提供方便。它既满足时空对称性，也符合理论本身所要求的对称性。它被称作维度正规化（dimensional regularization）。这种想法，即使是从量子理论的角度来进行评价，乍看上去也是非常奇特的。

一些方法会导致无穷大的出现，如果把理论建立在一个一切都存在于一条绳子里或一张桌面上的世界上，这种错误就不会出现了。也就是说，在一个总维度数是2或3（而不是4）的“玩具”理论①中，有一些量的总和不是无穷大，而是有限的数值。这给我们带来了一些启示：既然求和在不同的维度下能进行得更好，我们就把维度从4变到一个比4稍微小一点的数值。整个计算过程把维数当作一个参数，但如果最终把维度数设置为真实值

① 译者注：指那些虚构的但对真实物理研究有帮助的理论。

4，那么答案也会再次变为无穷大。

维度正规化的奇特之处在于，维度的数值不必是整数。想象一个三维世界的情形是比较简单的，根据经验想象一个四维世界也不是特别难。但是比 4 稍微小一点的维度数表示什么意思呢？很幸运，物理学家们不需要考虑它的真正含义，把维度看做一个连续的量只是一种处理问题的方法，这决不会出现在任何真实的实验结果中。

运用维度正规化的重要结果，就是将令人头痛的无穷部分从计算中分离、孤立出来。这样就容易处理了，而爆炸性的发散本身会呈现为某个特殊的、容易理解的形式。

这种大家或许都不怎么熟悉的改变维数的方法，在理论物理中实际上已经有了漫长的历史。尤其是在 20 世纪 20 年代，波兰物理学家西奥多·卡鲁扎（Theodor Kaluza）和瑞典物理学家奥斯卡·克莱恩（Oskar Klein）发展了一套将广义相对论和电磁学相结合的理论，其中引入了一个额外的维度——也就是总共包含 5 个维度。在阿根廷的拉普拉塔国立大学，卡洛斯·博利尼（Carlos Bollini）和胡安·吉亚姆比亚吉（Juan Giambiagi）在卡鲁扎和克莱恩模型的基础上，尝试研究跑动图（errant diagrams）中的无限大问题，讨论了奇数维中那些出现无穷大的图有些什么性质。他们不久就意识到，维度本身就可以作为处理无穷大的工具。维度正规化就这样诞生了，但在此之前，论文的发表经历了非常曲折的过程。博利尼和吉亚姆比亚吉的第一篇论文，在 1971 年 10 月寄到了《Physics Letters》杂志。但是这篇文章的审稿人并不关心高维理论，因为现实世界是四维的。他们讨论维度正规化的第二篇论文在 1972 年 2 月 8 日寄到了另一家杂志。几天后，第三家杂志社收到了另一篇有关维度正规化的独立文章，它是由荷兰的乌德勒支的杰拉德·特霍夫特（Gerard't Hooft）和马丁内斯·维尔特曼（Martinus Veltman）寄来的。然而特霍夫特-维

尔特曼的论文比从阿根廷寄来的论文刊登得更早，因此人们曾一度认为维度正规化的创始人仅仅是特霍夫特和维尔特曼两人。博利尼曾发表评论说："从阿根廷拉普拉塔寄出的论文到达编辑桌子上的时间，要比从欧洲任何城市寄出的都要长。"

运用维度正规化的优点十分显著，所以我们不妨先把那些非整数维度的物理意义搁置到一边。即使应用了维度正规化方法之后，费曼图本身的数学表述基本上也没有影响，费曼图的形式仍然可以辨认。另外，维度正规化还保留了所有可能的对称性。最后需要说明的是，在高能物理中常用到一种简化方法，就是将一些参与实验的粒子质量设为零，这是因为它们的质量与加速器的巨大能量相比显得非常小。但在理论上，这会导致一些问题的出现，因为在计算中将质量设为零会引起红外发散的产生。运用维度正规化，可以通过把维度增加到比 4 多一点点，轻易地解决这个问题。

在上面的讨论中，我们忽略了一个细节，但实际上不仅仅是细节，它会产生相当大的影响。在任何数学方程中，能够相加或相减的两项必须具有相同的维度。一个用米表示的距离加上另一个用米表示的距离，就能得到一个总的距离，并且也用米来表示。但是一个用米表示的距离加上一个用立方米表示的体积，结果是没有意义的。

QED 理论和 QCD 理论中的方程当然也具有确定的维度。但是如果像在维度正规化中的做法那样，比如将 QED 理论中的时空维度乱改，那么电子和光子之间相互作用的情况就会变得一团糟。

实际上在 QED 理论中，混乱的部分牵扯到耦合常数，当时空维度改变时，耦合常数的维度会变得不可思议。解决这个问题的方法是引入一种补偿因子，当然这是凭空引入的，因子的维度用来同耦合常数中的维度抵消。当把两项相乘时，其积的维数就

是普通维数了，这是可以接受的。

接下来我们会发现，这个新的参数在无限费曼图表达式的有限部分中一直存在。这看起来是利用维度正规化伎俩的惩罚，我们对这个引进的参数一无所知，我们只知道引进它只是为了解决一个技术问题，为了凑出一个正确的维度来，但最终却没有办法将其消除。

这会是一个致命的漏洞吗？恰恰相反，实际上我们并没有指定这个参数的值，它可以取任意值，这会导致一些令人吃惊的新物理现象。因此，有必要给这个参数起一个名字，并且找一个符号来表示它。它被称作重正化尺度（normalization scale），用 μ 来表示。实际上，这个表面上很神秘的量，与动量截断等效，只是动量截断在正规化方案中的表现形式。μ 这个量划分了“短程”与“长程”之间的界限。人们认为长程包含了我们关心的所有客体。排除短程实际上承认了我们并不知道短距离内的实际情况，但我们假设这不会造成任何影响，起码暂时是这样。

寻求“抵消法”

这样，维度正规化就提供了一套重新将费曼图的圈发散简化的方法，将其变为无限部分加上一个有用的有限部分。现在再看看重正化本身。实际情况似乎相当令人沮丧，因为在 QED 理论、QCD 理论以及电弱相互作用理论中，存在着无穷多的费曼图，几乎所有费曼图都包含着圈以及圈中无法控制的动量求和。实际上情况并没有这么严峻，理论物理学家们已经经受住了挑战。

我们再来研究一下 QED 理论，因为 QED 理论集中表现了问题，又没有那么多烦人的细节，因此我们只需要研究 5 个基本圈图就够了，见图 3.8。其他所有的费曼图，即使包括最复杂的圈，也可以通过这 5 个图来解决。

例如电子自能量图，见图 3.8 的图（a），维度正规化可以很

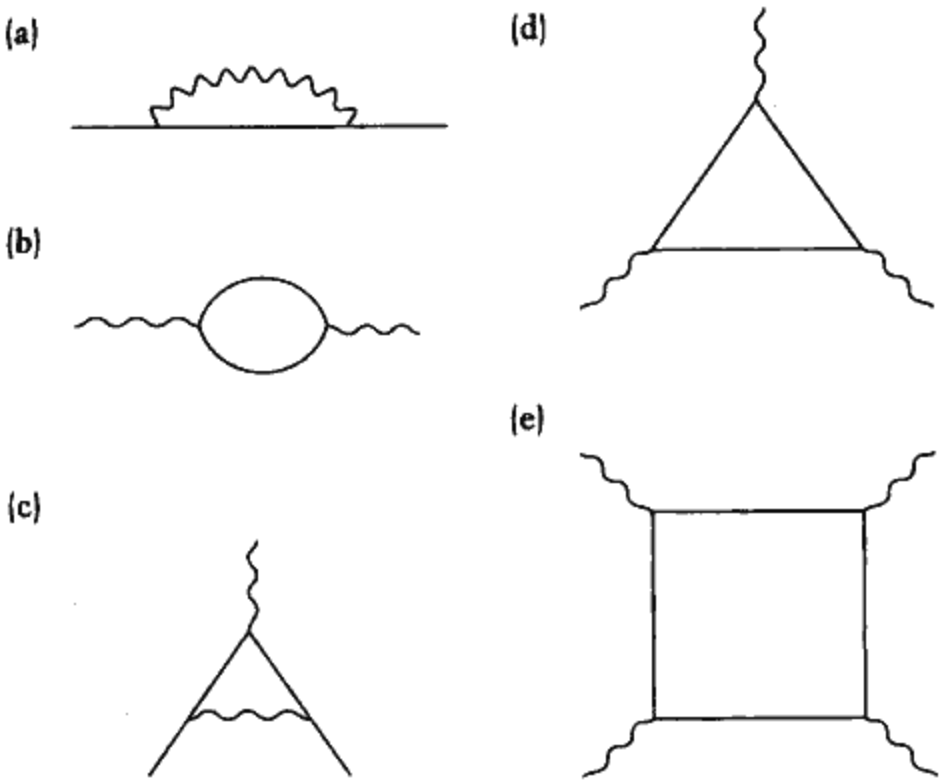


图3.8 QED理论中的基本圈图：图（a）是电子的自能量图；图（b）是光子自能量图，或真空极化图；图（c）是电子－光子顶点图；图（d）是三光子顶点图；图（e）是光子－光子散射图。实际上，图（d）不起作用，图（e）是有限的。QED理论中所有发散图都可以通过图（a）、图（b）和图（c）来处理

明确地划分出一个无穷大部分。重正化在QED的主方程中，额外加上一个简单部分，然后将主方程放到微扰论中处理。新加的部分也会得出一个无穷大的结果，与从电子自能图得到的无穷大部分数学形式相同但符号相反，因此这两部分可以互相抵消。主方程中新加的部分被称作“抵消项”（counter-term），这是因为它在效果上抵消了结果中的无穷大。在光子的自能量图和顶点图中，还需要把另外两个抵消项加到主方程中，用来抵消无穷大。

魔幻之旅从此开始了，尽管从这开始，看起来有点运气的因

素在里面。之所以说运气好，就在于要处理三光子顶点图和光子-光子散射图根本不需要进行任何操作，分别见图 3.8 的图（d）和图（e）。三光子顶点图，总可以同一个电子圈方向相反的（方向在图中没有标出来）三光子顶点图一同约掉，因此可以不出现在图中。最后一张图是光子-光子散射的“正方”图，由于这个方形中包含了那么多光子，结果是有限的。

但这并不是故事的结局，因为在一些实际物理过程的计算中，下一步需要包括更多的图，并且包含更多的圈。这些微扰级数中次数更高的项会带来什么后果？这些图会包含更多错综复杂的圈，这使得在单圈图中取得重大突破的抵消项毫无用处。

接下来要讨论的是具有两个圈的费曼图。例如有两个圈的真空极化过程（图 3.9），由于具有封闭的圈，这 3 个图都会出现无穷大的求和。这些圈相当复杂，算起来也许会出大麻烦。

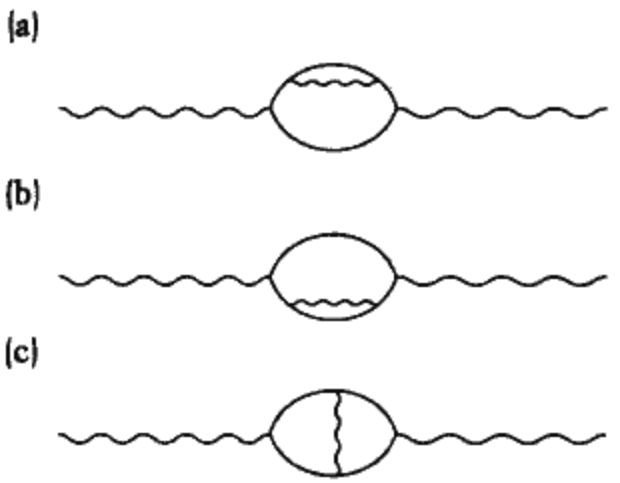


图 3.9 光子自能量的两圈图

但实际并不是这样，这正是神奇所在。回想一下加到主方程上的抵消项，它将由真空极化单圈图引起的无穷大抵消掉。这种抵消是在主方程、抵消项等都通过微扰论处理后才出现的。再进行一次这种工作，抵消项会对高次项产生什么效果？令人惊奇的

是，所有的费曼图（包括两圈图）都包含了由抵消项而产生的额外部分，将各种各样的项不可思议地抵消掉。剩下的项，在适当的精度上都是有意义的，只有最后一项是发散的，但这一项可以由下一个精度级别的抵消项抵消掉。

这就是问题的核心。对于像 QED 理论、QCD 理论以及电弱相互作用理论这样的重正化场理论，把一个很小的简单抵消项加到主方程中，就足以抵消掉所有次项中的无穷大项。这就同变魔术一般。

它之所以神奇就在于，一般要抵消新出现的下一级发散就要加入新的抵消项。但是根本不需要这么做，在 QED 理论中，原来引入的三项已经足够了。换一种角度来看待这个问题，在 QED 以及其他一些重正化理论当中，随着在微扰级数中计算越来越多的项，发散的部分同简单的圈图中出现的微扰项相比，既不会多一些也不会少一些。复杂的多圈图在本质上都是一样的。

还存在根本不包含发散的理论，在这种情形下一切都用有限的数值表示。其中主要的理论基于费米子和玻色子之间的奇怪对称性，称作超对称（supersymmetry）。尽管没有直接的实验证据支持超对称思想，但由于它是用有限值来表示的，所以在理论物理学家中非常受欢迎。

在一些玩具理论中，更高次数的图最终不再含有发散部分，也就是说理论包含有确定数量的发散图。这些理论是超重正化的，但是描述真实世界的理论都不具有这个性质。在其他玩具理论以及量子引力论中，发散会随着多圈图次数的增加而更加混乱。第一次引入的抵消项不足以抵消掉发散，在每一级新发散项出现时必须引入新的抵消项。这一系列无穷无尽的抵消项就是非重正化理论的标志。

先不说选择抵消项的技巧，加到主方程上那一部分的任务，就是消掉无穷大的部分。这里就要用到维度正规化方法，因为它

将烦人的发散部分归结为一个简单的附加贡献。因此寻找一个有用的抵消项就变得很容易了。尽管在选择抵消项时还有严格的选择范围，但无论怎样，物理学家们都会得到同样的结论——详细内容我们稍候再讨论。

这种方法还有两个讨厌的遗留问题，除了有一点头痛之外，或许还有一点被欺骗的感觉。例如，重正化思想会把像电荷、质量这样的物理量重新解释成什么样子？在 QED 或 QCD 理论的主方程中加上一部分，又期望它的物理结论没有变化，这会有什么意义？

实际上，那些部分并不是真的加到了主方程上；而更像是，先把主方程重写，然后对其进行谨慎和奇怪的解释。一开始，主方程是用纯的、不可测量的量表示；在电动力学中主方程包括纯电荷和电子质量。这些“裸露的”物理量被认为是无穷大的。这种解释会带来一些麻烦，因为纯物理量在实验中是无法观测到的。因此，没有经过重正化的裸物理量首先是无穷大的，有圈的图会引起发散，接着把这些量抵消掉，就剩下了有限的重正化物理量。重正化工作的目的就是要实现这一点。

一套重正化理论只能有确定数量的发散项，这样才能处理所有可能的发散。在电动力学中有 3 个基本的发散图，与 3 个抵消项以及 3 个无穷大的常量相对应。电子的自能图将电子的裸质量转变成物理质量。顶点图联系着裸电荷与物理电荷；光子的自能图可以将裸电场强度转化为物理电场强度。光子本身的质量在重正化前后都是零。

要在一个重正化理论中做出物理上的推断，实验物理学家需要进行足够多的测量以确定理论中的参数。例如在电动力学中，就意味着要测量电子质量与电荷。这样，理论才能用来预言一些其他过程的结果。

这让我们对非重正化的意义认识得更加深刻。在一个没有被

重正化的理论中，微扰展开的每一个新项都要加上新的抵消项。这不仅表示要确定所有的结果就要做无穷多次试验，还表示存在无穷多数量的“物理”参数。这说明了一个核心问题，就是像量子引力论这种非重正化理论不具备预言物理现象的能力。像 QED 和 QCD 以及电弱相互作用理论这样的重正化理论，所有的发散都可以通过理论中物理量的无限次重正化消除掉。

“骗人的办法”实际上是对的

理论物理学家们可以计算一整套费曼图，但对于高次项——就表示要计算更多的圈和更多的图，计算量是相当大的。想要检验重正化方法能否行得通，仅仅按部就班地计算越来越复杂的费曼图也不是一个现实可行的方法。

此外，通常人们也不是这样用数学处理问题的。也就是说，数学证明并不是通过列举所有可能的情况而完成的。数学家们首先要把所有的情况总结为一个普遍适用的表述形式。在重正化中，这就意味着需要某些人坐下来证明出一个普适的结论：尽管复杂的费曼图越来越多，但用来抵消发散的各项仍是可以合并的。

1949 年，戴森在普林斯顿高等研究院工作时，恰恰成功地证明了这一点。在一篇经典论文中，戴森建立了一套分析费曼图的术语，并且证明了量子电动力学的重正化对所有次项都成立，也就是说，如果一个低次项符合一套重正化理论，那么所有可能的费曼图都符合这个理论。他的证明是用数学归纳法完成的。他首先假设重正化对某一次数的费曼图成立，接下来证明该过程对比它高一次的费曼图也成立。

所有带有圈的费曼图都可以用很小数量的基本费曼图（在 QED 理论中，这个数量是 3）来分析，戴森就是靠这一点完成证明的。但是有许多图具有重叠的圈，这些图分析起来特别难。人

们必须理解这些互相重叠的圈，来填补戴森工作的空白，这个任务落到了剑桥的研究生阿卜杜斯·萨拉姆（Abdus Salam）的肩上，他在1950年理清了重叠圈的问题。戴森论文中的另一个假设被纽约哥伦比亚大学一位很有前途的物理学家——史蒂芬·温伯格（Steven Weinberg，图3.10）所证实。温伯格在1960年发表的“收敛定理”中说：一般地，经过重正化，求和实际上很容易计算，即便费曼图中含有很复杂的圈，计算也很简单。

图3.10 1992年的史蒂芬·温伯格。温伯格因提出电弱理论以及现在所谓的标准模型（包括强力、弱力和电磁相互作用的理论）而获诺贝尔奖，也因此声名显赫。图片来源：刘易斯·温伯格

具有讽刺意味的是，后来温伯格和萨拉姆对他们自己的弱力理论中的重正化问题感觉不满，而他们正是因为这个理论声名卓著的。在美国物理学家谢尔登·格拉肖（Sheldon Glashow）工作的基础上，温伯格在1967年，萨拉姆在1968年分别构想出一套有欠缺的理论，后来特霍夫特和威尔特曼证明这就是重正化理论。

同一年（1971年），特霍夫特和他的导师威尔特曼又证明，所有QCD类型的理论也同样是重正化的。威尔特曼和特霍夫特（图3.11和图3.12）因他们在驯服无穷大问题方面的工作而获得

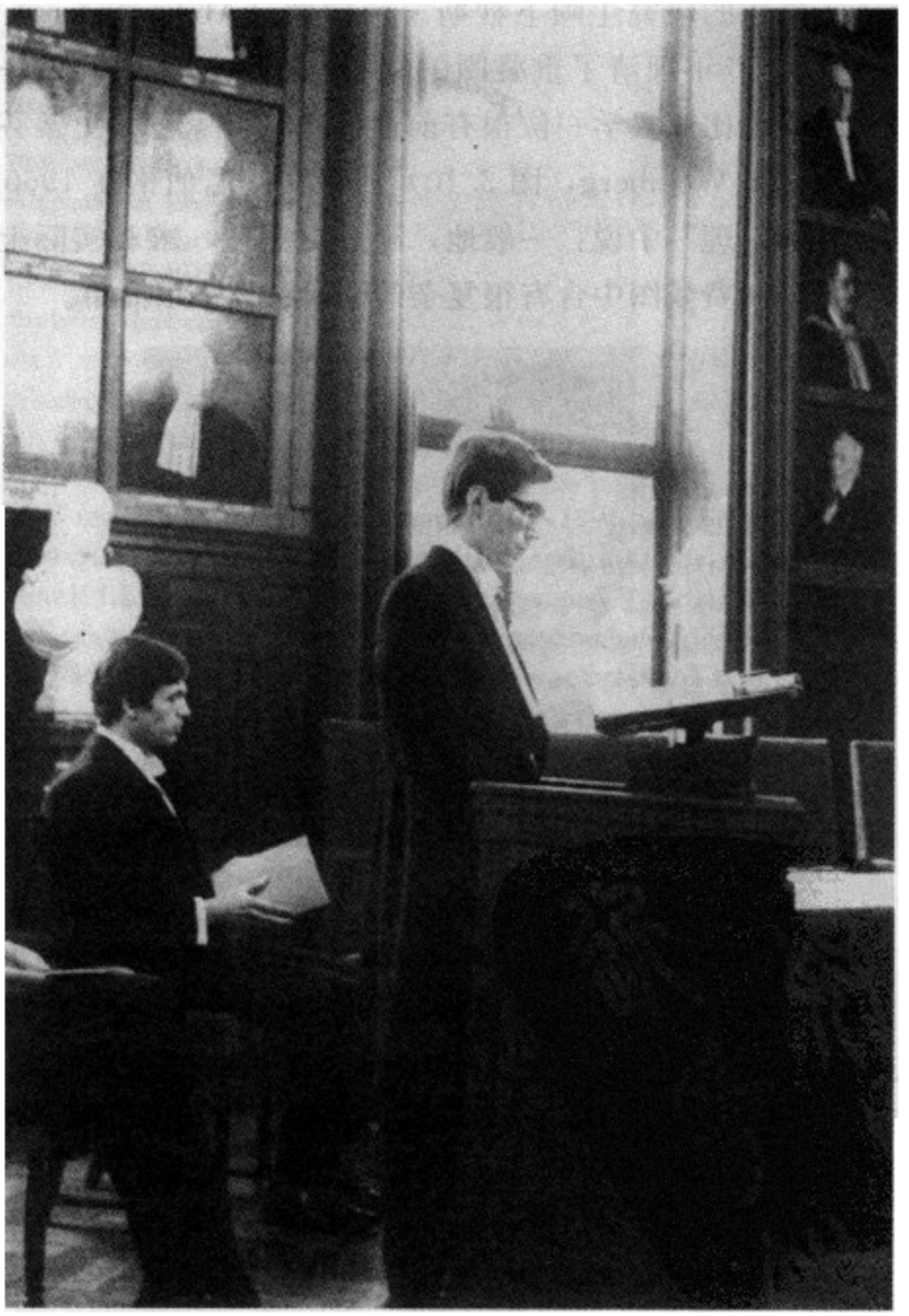


图 3.11　着装优雅的重正化：1972 年 5 月 1 日，杰拉德·特霍夫特正站在乌德勒支大学的讲台上进行研究生论文答辩。图片来源：范·德·斯鲁特，得到杰拉德·特霍夫特的授权

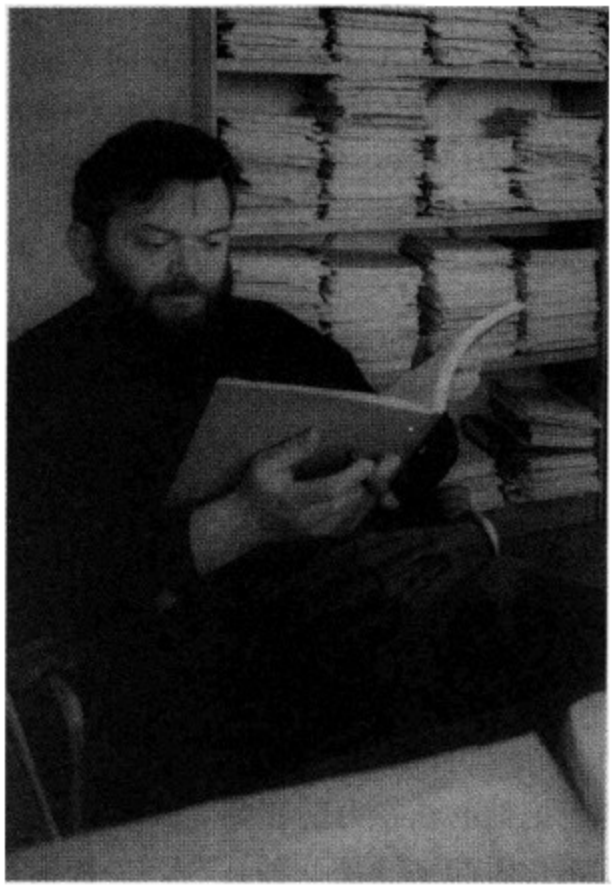


图 3.12　马丁内斯·威尔特曼于 1973 年在约翰·贝尔位于 CERN 的办公室里。威尔特曼深信出现在基本作用力理论中的无穷大能够被驯服，这鞭策着他同他的学生杰拉德·特霍夫特穿越重正化的迷宫寻找道路。他们两人由于“阐明了物理中电弱相互作用的量子结构”而分享了 1999 年的诺贝尔奖。图片来源：CERN

了 1999 年诺贝尔物理学奖。早在 20 世纪 30 年代，发散问题就已经让迪拉克（Dirac）、罗伯特·奥本海默（J. Robert Oppenheimer）、海森伯（Heisenberg）以及其他物理学家头疼了，将这个问题转化为对现代粒子物理有用的形式，耗费了全世界最好的一些理论物理学家 40 多年的时间和巨大的努力。

如前所述，理论物理学家可以随意选择重正化方法，就选取

抵消项的自由程度方面，我们再总结一下。选择范围依赖于将主方程分为物理上有意义的部分和无穷部分的方法。然而最终的结果不依赖于任何这种选择，因为这种选择是任意的。这种不依赖于选择的性质是一种不变性，这种不变性揭示出一种对称性。这种对称性在数学上表述为重正化群——该群在 QCD 理论中有特殊用途。

重正化的尾声

人们认为，世界上有 4 个人能解决电动力学中的发散问题，他们对重正化理论发展到今天这个样子作出了很大贡献，但他们没有一个人是重正化的忠实信仰者。费曼不喜欢重正化，这从他把重正化评价为“脑筋不正常的过程”就能看出来。朱利安·施温格（Julian Schwinger）在这个问题周围徘徊了好多年，朝永振一郎（Sinitiro Tomonaga）把它看做是一个临时性的解决方案。而沮丧的戴森也把他的注意力转移到了其他方面。

最初，重正化看起来像是一种技巧，一种实用主义催生的胡搅蛮缠的理论，是为了从看起来很优美但难以处理的理论中得到有意义的结果。从 QED 理论的成功经验来看，作为一个规范化原则，重正化非常重要，它是任何一个实际可行的理论都必须通过的检验。同样，描述弱力的电弱理论以及描述强相互作用的量子色动力学都没有捷径可走。

但最近一段时间，受到建立一套重正化量子引力理论想法的启发，理论物理学家们把他们的注意力重新投入到了重正化上。当然也存在其他情况的可能性：或许在适当的条件下，坚持一套即使不符合重正化要求，但看起来又很漂亮的理论也是有意义的。有些人认为，我们应该学会适应那些不完备的、但能帮助我们解决感兴趣的问题的理论。

在我们的宇宙中，存在不同的长度尺度，这反映在存在不同

的现象以及描述这些现象的各种理论中。这种长度的范围，一端延伸到亚原子世界，然后是原子尺度、分子尺度，然后是“日常固体”和恒星，接下来是星系，最终是整个宇宙。让原子物理学家们着迷的那些结论，对原子核内部的细节无能为力。建筑桥梁的工程师对钢筋内部的原子细节不感兴趣，就如同天文学家们对大桥的精巧结构毫无兴趣一样。

每一个研究组，都会有他们自己处理问题的尺度范围，也会有在这个范围内对世界运行方式的描述。尽管一位原子物理学家感兴趣的原子存在于星系中，但他并不需要了解星系尺度上的问题，他可以把一个星系看做是无穷大的东西。一位工程师也不需要考虑原子大小上的问题，他可以仅仅把原子大小当作是零，而把大桥钢筋看作是一种真正连续的材料。

实际上这是一种伟大的思想——仅仅把注意力集中到感兴趣的长度范围内，而把更小尺度中的东西看做是零，将更大尺度上的东西看做是无穷大。结果会得到一个“有效理论”。实际上比较小的量并不是真的零，比较大的量也不是真的无穷大，如果需要的话，可以以该理论为起点，用微扰论的方式将这种偏差计算进去。

当化学家们讨论原子通过共享或提供电子从而结合成分子，其实就是在运用“有效理论”。桥梁工程师用到的力学是另一种有效理论。

在量子场论中，有效理论方法通常是指，在非常短的长度范围内长期观测所发生的事情。设想有一套量子场论，它对于某种类型的过程以及某一个能量范围内的问题描述得相当好。很容易想象，在比它高得多的能量下，也就是指在短得多的距离上，会出现一些原来理论中没包含的新现象；就像一位工程师如果用一台强大的显微镜观察钢筋材料时，就会发现他原来认为连续平滑的材料是错综复杂的。有效量子场论的目的是，构建一套能够让

理论物理学家们在某几个能量范围内处理问题的理论，而不必了解所有的内容——即不必了解自然界在高能范围内的所有规律。

有效量子场论的思想是由温伯格在 1979 年引入的，并且逐渐得到很多有力的支持，它看上去非常像反向跟踪。但是毫不留情的重正化要求使这种优秀的思想岌岌可危。有一个经典的故事，可以说明理论物理学家们如何支持他们的这种修正方法。那还是在 1933 年，恩里科·费米（Enrico Fermi）提出了一套有关中子通过弱作用辐射衰变为质子的理论。假如把它应用到 μ 子衰变中，费米的理论就表示，在 μ 子消失的地方会出现一个电子、一个中微子和一个反中微子。费米的理论能给出一些很有意义的结果。例如，很容易就可以算出衰变 μ 子的寿命，并且计算结果与实验数据十分吻合。这很不错，费米的理论不可能彻底错误，但肯定是这样吗？

对于低能量情况下的简单计算，费米的理论看上去很管用。但是如果和更精确的实验对比一下，或者用它计算一下更高能量下的散射截面，结果就出错了。费米的理论是非重正化的——他的理论中丢失了传递相互作用的 W 粒子和 Z 粒子，这是至关重要的，但在费米提出这套理论的时候，这两种粒子还没有被发现。因此以重正化的标准来衡量，费米的理论是不成功的。但是现在理论物理学家们还是会把它称为“有效”量子场论——它是一种在某一特定能量之下，具有很好效果的非重正化理论。这个特定能量一定要比 W 粒子和 Z 粒子的能量小。所以我们并不是要把费米的理论扔掉，而是根据有效理论的应用范围，在它有效的情况下使用它，并且还要给尚未发现的新物理留下一定的空间。当然，在弱相互作用理论中，物理学家们现在已经知道其中的新物理内容是什么了，那就是传递弱力的 W 粒子和 Z 粒子，随之诞生的新理论就是电弱相互作用理论。

那么像 QED 这样成功的理论，今后会向什么方向发展呢？

它是有史以来最成功的理论，并且也是重正化的，这就表示所有的量子场论要想成功就必须是重正化的。但根据有效场论的观点，迄今为止，QED 理论的成功只不过隐瞒了一个事实：那就是物理学家们对它的检验还不够充分。这为实验物理学家们测量电子和 μ 子的磁矩带来了新的动力，因为通常人们把磁矩看做是理论能够精确处理的物理量。现在，这些测量还对质量和长度范围做出了限制，在该限制下会产生一些超出传统 QED 理论范畴的新物理过程。在将电子和 μ 子的磁矩计算到小数点后几位的精度下，光与物质的相互作用可以成功地用一套理论来描述，这套理论的主要机制就是通常的 QED 理论，一套可以重正化的理论。如果再加入描述光与物质相互作用的任何其他机制，最终都意味着加入了不可重正化的项，这些项只有在极高能量以及极短距离上才会出现。一般在实际可用的理论中，都会包括重正化部分和非重正化的部分——重正化已经丧失了作为基本相互作用理论所必须满足的基本要求的地位，这对于研究引力的理论物理学家们来说是一个好消息。放宽了重正化的要求后，取而代之的是，有效量子场论的结构需要满足一定的限制。有效场论是一次概念上的突破。

因此，现在 QCD 理论和电弱相互作用理论都应看做是不完备的理论。这些理论会在某个特定能量之下（对应着很小的距离尺度）适用，新的理论要么从计算与精确测量的细微差别中诞生，要么会从巨大能量的碰撞中诞生。在理论物理学家们的写字板上，接下来出现的内容仍然是一片空白。但从现在开始，我们要和“传统”的 QCD 理论一起，开始一段新的旅程。

加入对称性

这是理论拼图的最后一块。量子场论提供了一个将世界的量

子性和相对性结合起来的框架。时空变换下（包括转动）的对称性被植入了这个框架中，但是还落掉了基本量子成员的对称性。如果把这种对称性也通过某种形式包括进来，结果就形成了规范理论（gauge theory）。

在这项游戏中，最关键的成员就是相位。波动的相位就是指一个表示波动进程在周期内的角度，两列波的相位不同就说明它们不同步。规范理论与通过相位表示的对称性有关，它依赖于外部环境和量子系统相位之间的特殊关系。直到最近，人们才通过实验清楚地展示了这种关系。

重谈双缝实验

我们再来谈一下双缝实验。通过双缝的光，由于干涉作用会产生明暗相间的条形图样。在效果上，双缝就是两个分开的步调一致的光源，每列光波都可以用振幅和相位来表示。在探测屏上两个波峰（或两个波谷）重叠的地方，振幅会变为原来的 2 倍，亮度增加为原来的 4 倍。中央亮纹对应着两个步调一致且到双缝距离相等的波峰（或波谷）叠加——对于中央亮纹，从两条狭缝过来的两列光波没有相位差。两边的一级亮纹对应着从一条狭缝射来的光比从另一条狭缝射来的光整整提前或拖后一个周期，这是由两条狭缝到探测屏的距离差造成的。再向外看还会有另一对亮纹，每边各有一条，这对亮纹对应着从一条狭缝射来的光比从另一条狭缝射来的光提前或拖后两个完整周期。屏上每一条亮纹所对应的光波波峰之间都没有相位差，或者相位差是周期的整数倍。而暗纹对应着相位差是半个周期、一个半周期等。

到目前为止，我们除了把前面章节中出现过的实验又做了一次详细描述外，再没有什么新的发现。现在，我们要把某些东西变一下。如果把一个“光减速器”放置在其中的一条狭缝后面，光减速器除了让通过它的光变慢之外没有其他的影响，那么现在

结果会有什么变化？屏上仍然会产生干涉条纹，但由于从其中一条狭缝射来的光花的时间要稍微长一些，因此所有的相位差都会有一些改变。结果造成整个图样向左或向右偏移，但图样本身并没有任何改变。由对称性可知，如果把光减速器放到另一个狭缝后面，图样就会向相反的方向偏移。顺便说一下，光减速器并没有进行任何把其中一条光线和另一条光线区别开来的操作，因此它也不会破坏干涉图样的光子从那条路径经过的信息。

如果用光减速器把两条狭缝都遮住，会发生什么情况？其结果看起来和原来完全一样：图样不会向任何一边偏移。这是因为干涉依赖于相位差，而不是相位本身。因此从两条狭缝中射出的减速光的相位一同变化，但是对相位差没有影响，因此整个图样也不会表现出任何变化。相位差决定了明暗条纹的位置以及它们的相对亮度，波动的振幅决定了干涉条纹的总亮度。在现实中没有办法观测到相位本身。

用光做的基本双缝试验很简单，在一个高中实验室中就可以完成。相反，一个用单个电子通过机械双缝而实现的真实双缝试验从来没有人做出来过。这是因为它太难了，所需的仪器小得不可思议。但是人们通过另一种方式完成了电子的双缝试验，在两块导体板中间夹上一根带正电的细金属丝，板与金属丝之间的空间就形成两道“狭缝”。金属丝与导体板之间的电场将电子集中到中心区域，就像光学中的双棱镜效应一样。在 1989 年，日立东京研究实验室一个由外村彰（Akira Tonomura）领导的小组，将一个电子双棱镜插入到一个被称作场致发射显微镜的仪器中，然后让一个电子穿过该装置，并在装置后面安放一台对位置十分敏感的电子计数器。每次只有一个电子到达计数器，逐渐就会在计数器上出现预期的独特干涉条纹。

如果外村彰和他的同事将一个电子减速装置放到其中一条狭缝的后面，他们一定也会观测到图样偏移的效应。如果在两条狭

缝后面都放置电子减速装置，图样将没有任何变化。在该实验中，观察到电子的相位差是可能的，但是并不能观测到电子的绝对相位，绝对相位在任何其他实验中也不可能被观测到。对于光子和电子来说都一样——改变从双缝射出来两束粒子中一束的相位，最终得到的干涉条纹会以大家容易理解的方式移动。

我们可以用沿一条笔直的马路开车的情形类比，实验中感受到的是相位差，而不是实际相位。就像一个自身不断做着周期性重复的波一样，汽车的轮胎也在一圈一圈地转着。四个轮胎转动的步调一致，司机根本不会知道每只轮胎处在它们一圈中的什么位置——例如，他不会知道任何一个轮毂盖上的制造商标志什么时候恰好朝上。但如果车经过一个在副驾驶一侧的水坑时，将会造成车的那一边减速，并且立刻使副驾驶一侧的轮胎减慢速度。如果恰巧四个轮毂盖上的制造商标志在车经过水沟前都是冲着一个方向的，那么经过水沟后就不是这样了。水沟“减速器”使汽车两边的车轮产生相位差。司机会注意到，由于没有纠正措施，他会转弯开到旁边的沟里去。

侦探小说迷们或许能够“嗅”到接下来的前进方向。在其中一条狭缝上放置减速装置会产生明显的效果，而在两条狭缝上都放置减速装置就没有任何效果。这种情况的解释是：实验遵从在电子波相位整体变化情况下的不变性。量子理论的思想也是这样的，实验对相位差很敏感，而对绝对相位没有反应。如果任何情况与以上思想不符，那么所有书本上的内容都要重新改写，因为这是量子理论的核心性质。相位不能被直接观测到，因此物理效应都是相位整体变化的不变量，这就意味着存在与相位对称性相关的守恒定律。

还有另一个谁也没做过的实验。没有人完成这个实验是因为没有人知道进行这个实验的方法，尽管早在 1959 年就有人把这个实验提出来了。这是电子双缝实验的另一个版本，只是这一次

放置在每条狭缝后面的是一个中空的金属管，图3.13是该实验的示意图。实验的思想是让单个的电子射向双缝，等一段时间，让电子刚刚进入金属管，然后在金属管上加一个电脉冲，并在电子运动到金属管出口之前将电脉冲停止。

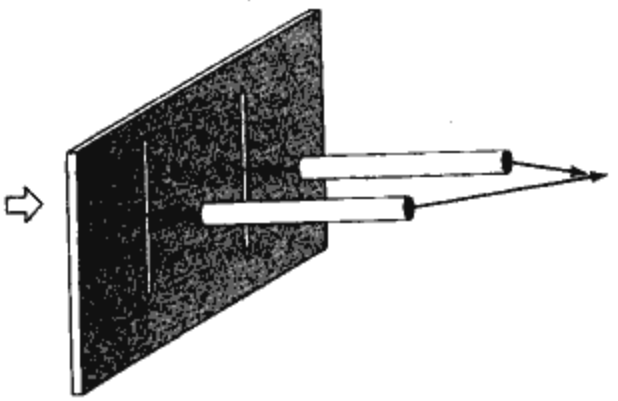


图3.13 用两根带电金属管演示的阿哈罗诺夫－玻姆效应。电子从左边进入，双缝中每条缝的后方都有一根可以充电的金属管

与电源相连的一个封闭的导体内部没有电场，即使外边有电场也是这样。因此在以上所讲的电子双缝实验中，金属管内部没有电场，当电子离开金属管时所加的电脉冲也停止了。因此人们可能会以为，由于并没有电场施加于电子上，因此不会有什么现象产生。接着人们会理所当然地以为即使在两根金属管上加不同的电脉冲，所得的干涉条纹也不会发生改变。当然这是一个经典物理学家的预言，因为根据经典物理理论，一个电荷只能受到电场或磁场的影响。

当完成这个实验之后，实验物理学家们对他们所观测到的现象会毫不质疑。当施加的电脉冲不同，干涉图样也会变化。当施加的脉冲相同，干涉图样就不会移动。他们知道除了电场之外，还存在其他电效应会影响干涉条纹。

一次实验杰作

物理学家们之所以对该实验的结果如此肯定，是因为这种效应是另一个已经从实验中观测到的现象的不同形式。图 3.13 所示的实验，实际上是阿哈罗诺夫-玻姆效应（Aharonov-Bohm effect）的一个版本，该效应是由亚基尔·阿哈罗诺夫（Yakir Aharonov）和大卫·玻姆（David Bohm）在 1959 年提出来的，当时他们都在英国的布里斯托尔大学。阿哈罗诺夫-玻姆效应的另一个版本已经得到了实验验证，图 3.14 所示的是该实验的示意图。这仍是一个电子的双缝实验，但这一次是将一个极细的电线圈（螺线管）放到双缝的后面，把它放置在两条缝之间并与双缝平行。当螺线管中通上电流后，螺线管内部会产生磁场，但管外却没有——这是螺线管一个众所周知的性质。因此当电子从双缝到达探测器的途中，会从螺线管旁经过，它经过的是一个不存

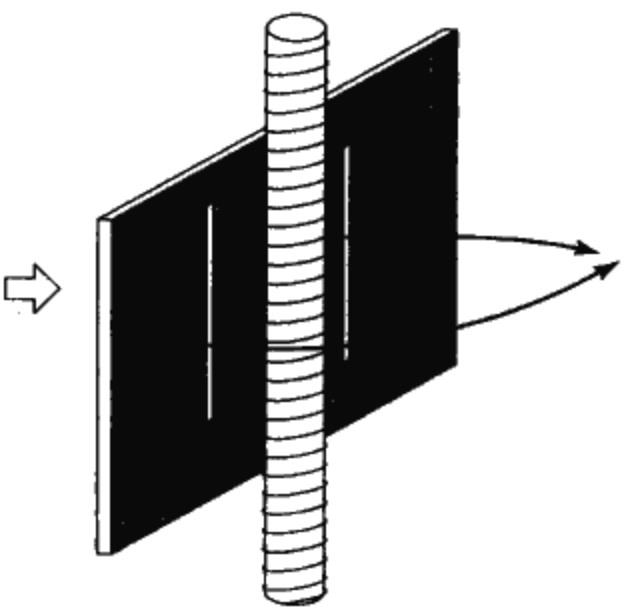


图 3.14　阿哈罗诺夫-玻姆效应。电子从左边入射。双缝的后方是一根螺线管。螺线管内存在磁场，但螺线管外部却没有

在磁场的区域。然而当线圈中有电流，电子的干涉图样就会发生变化。经典理论再一次失效了，由于不存在磁场，电子是没有理由以任何方式做出反应的。显然，除了磁场之外，还有别的什么东西在这里起作用。

在每个人都相信结果之前，实验需要重复很多次，但这个实验太难做了。在 1968 年一次精妙的实验完成之后，所有的怀疑都烟消云散了。这次实验也是由外村彰和他的小组完成的，实验中他们用到了一个直径只有几微米、外面包裹着一层超导体的圆形磁铁。这样，在这个实验中，长螺线管就被封闭到一个圆环中，磁场也被限制在圆形材料中。不管电子是从环中间的孔中穿过，还是从环四周通过，它都是自由的，最终会在屏上产生干涉。在小磁铁中通入电流后，尽管它周围的空间中不存在磁场，干涉条纹还是像预期那样发生了变化。

我们还可以用一种行为像小磁铁的、不带电的粒子做类似的实验，这是由阿哈罗诺夫和阿哈龙·卡什（Aharon Casher）在 1984 年首次提出的。与阿哈罗诺夫-玻姆效应的两种情况类似，该效应也有两种可能的版本，最近这两种实验都已经用中子实现了。

中子不带电荷，但是由于稍后就会说明的原因，它确实很像一个小磁铁。在 1989 年由澳大利亚墨尔本大学和美国密苏里大学联合完成了一项实验，实验人员用密苏里大学的科研核反应堆示范了其中一种版本。他们让磁化方向向上的中子通过带电金属丝的两侧。给定了磁化方向，中子虽然没有受到电场力，但干涉条纹还是像预期的那样发生了变化，这说明即使在没有力的情况下也存在一种影响。

在 1993 年，该研究小组运用与图 3.13 相似的设计，但用两根线圈替代了两根金属管，演示了最后一种可能的实验。中子的磁化方向与它们运动的方向相同，当中子经过线圈的时候，在线

圈上加以电脉冲，在线圈内部产生磁场。磁场没有对中子施加任何力，这是由于磁场均匀分布以及中子都朝着同一方向。但是尽管没有力的作用，实验物理学家们还是在干涉图样中看到了不可思议的变化。

这些表面上看起来十分神奇的影响实际上一点儿也不神秘，其实它们也有自己的名字。它们被称作“势”（potential）。在经典物理学中，很多人都了解“势”。比如说在电学中，两处电势的差别，有一个家喻户晓的名字——电压。电势在不同位置的变化就是电场，电场会对一个带电的粒子——如电子，施加我们经常遇到的力，这样电子就会感觉到它位于一个电势不断变化的区域。对于磁，情况稍微有点不同，但总的思想是一样的。在电场情形下，由电荷引起的势能是一个简单的数，它在各个方向上均匀地逐渐减小——电势能是一个标量。相应的电场呈辐射状地远离电荷或者向电荷汇聚，由于有方向，因此它是一个向量。螺线管周围存在的磁势，也随着到线圈距离的增加而逐渐消失，但磁势具有方向，它的方向是绕着线圈的。由于磁势有方向，因此它也是向量。有时磁势的图像用绕螺线管轴的同心圆来表示。圆圈上的箭头指示该势场的方向，它的方向同绕线圈的电流方向相同。

磁势围绕线圈的方式为理解阿哈罗诺夫-玻姆实验提供了一些线索。想象一下划独木舟，有一个人经过一个漩涡的一边，而另一个人经过漩涡的另一边；其中一个人是顺流，而另一个人是逆流。如果他们约好了在远处会合，顺流的人会先到达。对于电子经过螺线管来讲，从线圈的一边通过时它们是沿着磁势的方向的，而从另一边通过则逆着磁势方向。这就是造成干涉图样移动的相位差产生的原因，相位差是由于开启了线圈内部的磁场造成的。在线圈的一边，相位由于势的存在而增加；在线圈另一边，相位会减小。结果就造成了相位变化，相位变化依赖于势而不依

赖于场。这就是大自然在量子层次的运转方式——电子（以及其他带电粒子）所受到的电磁影响，是通过改变它们的量子波相位来实现的。

势：千变万化

站在我们头顶上方的高压线上的小鸟，证明了电势本身对实验结果并没有直接的影响。鸟儿位于高电势处，但是它身体的各个部分之间没有电势差，因而就不存在电场，因此鸟儿能够继续像平常一样做着自己的事情。但是如果鸟儿用它的喙去啄一条连着大地的电缆，那么鸟的喙和爪子之间将会存在电势差。在如此短的距离内，鸟身体内的电场强度将会相当大，鸟儿体内的机能将会遭到破坏。这只鸟可能会被烧焦。

像站在高压线上的小鸟一样，在极高电势下做的量子力学双缝实验（或许是在放置于同一根高压线上的实验室中做的实验）同样也会没有什么差别。实际上，在通着高压线的实验室中的实验者，或许正在用放置在双缝后的两根细导体管演示阿哈罗诺夫-玻姆实验，见图 3.13。他们在地面上的普通实验室中所观察到的现象，和在头顶高压线上的实验室中所观察到的将会一模一样；他们无法仅仅凭实验结果来判断他们处于哪一个实验室之中。

以上问题的关键就是，完成实验的电势改变了，但得到的结果没有任何变化，就像将两根金属管中的电势都改变同样的量后，实验结果也不会改变一样。实际上，在实验中只要与电子有关的电势在每个地方都改变相同的量，那么实验结果就会完全一样。每个地方的每个事物都改变相同的量，这种变化叫做全局变化（global change）。电势的全局变化使得鸟儿能够站在高压线上而不受到伤害，也使得双缝实验的结果不受影响。相反，电势的局域变化（local change，这里的局域变化是指每一处的变化不完

全相同）足以让鸟儿烤焦或者使双缝实验的结果发生改变。将图3.13中两根金属管的电势改变不同的量，就是一个简单的电势局域变化的例子。

在图3.13所示的阿哈罗诺夫-玻姆试验中，电势的全局变化肯定会给电子波带来影响，但是得到的干涉图样却不发生改变，特别是干涉图样的亮度没有变化，这说明打到屏上的电子数量没发生变化。这是一种不变性，与之对应存在着一种守恒定律。

实际上，通过仪器的粒子是不是电子并不是最重要的，重要的是它们带有电荷，只有这样它们才会具有电势。因此该守恒定律并不是关于电子的，而是关于它们所带电荷的。另一种证明的方法就是用其他带电粒子再做一遍完整的试验，结果也是一样的。

由于电势通过改变电子波的相位来影响电子，因此电势的改变要由相位的改变来补偿。电势的全局变化会引起电子波相位的全局变化；同样，电势的局域变化也会引起电子波相位的局域变化。那么上面提到的守恒定律应该这样表述：整体电荷守恒等价于系统在电子波相位全局变化下不变。

规范变换

德国数学家赫尔曼·威尔（Hermann Weyl）是第一个将电子波相位的全局变化同电荷守恒联系到一起的人。在1928年，他在没有阿哈罗诺夫-玻姆效应以及实验支持的条件下，另辟蹊径地得出了该结论。之前的几年，威尔一直试图统一引力和电磁力，在工作中他发明了一套比较时空中不同点之间长度标准的方法。他引入了一个表示长度尺度的名词，通常称之为“规范”(gauge)，这个词在所有粒子物理学家的脑海中刻下了永不磨灭的印记。“规范”一词最初是代表一些更具体的东西，比如铁路工程师用“规范”表示铁轨之间的间距。威尔最初的统一理论以

及赋予“规范”的长度尺度意义都已尘封在历史书中，存留下来的只有“规范”这个词语本身以及威尔所用方法的形式。

近来，人们常用到“规范变换”（gauge transformation）这个术语，例如用它来讨论电磁势问题，在经典电磁学中可以改变电磁势而保持电磁场不变。电磁势是不唯一的：你可以自由地定义它们，就像可以移动的球门柱。正是因为这种可以重新定义的特点，才导致了“规范变换”这个术语。最重要的是，这些变换可以保持麦克斯韦方程组的形式不变。

其实，电势的局域变化（例如像电荷相对于另一个电荷运动的情形）要满足麦克斯韦方程组，就必须对磁势也做相应的局域变化。在 QED 理论中，这两种势实际上是光子性质的两个方面。

在量子理论中，势的变化需要用相位变化来补偿，因此“规范变换”是用来表示量子场的相位变化的。这样，理论就会遵从这种相位变化的不变性。

整体电荷守恒可能与全局相位变换相关，但全局变换仅仅是整个故事的开始。无数实验证明，电荷是守恒的，但实际上其中隐藏了许多秘密；如果电荷只是整体守恒，那么就有可能有一个电荷在一个地方消失，而另一个电荷在别的地方出现。在这种情况下，所有电荷的总量是守恒的，但是“别的地方”可能离得特别遥远，甚至可能是在另一颗行星上。这不仅违背了狭义相对论（因为这就意味着超光速的传播），还表示电荷守恒是观测不到的，除非你可以在瞬间观测整个宇宙。这样，如果一个电荷在地球上消失了，人们可以在另一个遥远的行星上放置一个探测器等着探测电荷的出现，从而证明电荷守恒。实际上，如果电荷真的只满足全局守恒，电荷守恒就不可能在实验中真正被观测到。

一个比较合理的解决方案是认为电荷满足更严格的守恒，也就是局域守恒，这就意味着一个电荷湮灭了，就一定伴随着在附近有另一个相同电荷诞生。所谓“附近”就是要受狭义相对论的

约束，也就是两个过程不能以比光速更快的速度发生联系。如果满足电荷局域守恒，就意味着存在局域的规范变换。在量子理论中，这就意味着量子场的相位变化依赖于位置，这与相位全局变化相反，全局变化中各处的相位变化都是一样的。

因此就存在两种规范变换——局域变换和全局变换，有时也分别称作第一类规范变换和第二类规范变换。分辨这两种变换的一种方法是，想象一群在终点处的看台上看赛马比赛的人。当赛马接近终点时，所有人的头都会一致转向终点，看到底是哪匹马先闯线——在这个例子中，这是一种人头转动的全局变换，因为所有人的头都以统一方式变化。当最后一匹马通过终点之后，人们开始与邻座的人讨论赛马的结果，有些人开始寻找卖爆米花的人，还有的人在看自己的赛程表。这时还是存在转动变化，但在看台各处人头的转动各不相同，因此是一种局域变换。

许多实验以及电荷局域守恒的经验说明，电子-正电子量子场的相位变换也应该是局域的。如果强制施加这种变换，并且要求描述自由运动粒子的理论在整体上仍不发生改变，就会发生十分不同寻常的事情。

在电子-正电子场相位的局域变化中，保持对称性的唯一方法就是引入一个新的场。这个场的变化能够消除电子-正电子场相位的局域变化所带来的影响，否则局域变化就会破坏对称性。为了完成这种补偿，新引入的场必须以某种方式和电子场相联系——最一般的联系方式是耦合，而耦合强度在主方程中是用耦合常数来表示的。

由于大自然拥有一种不可思议的经济性，这种新引进的场（一般称作规范场）其实一点儿也不“新”：它就是光子。这种光子就是所谓“规范玻色子”的一个例子。规范玻色子是一种具有整数自旋（例如光子的自旋值是 1）的、传递力的粒子，它在规范理论中出现是为了保持局域规范的对称性，相应的耦合常数恰

好就是电子所带的电荷，这使得大自然的经济性体现得更加完美。

在一个包含自由电子和正电子的理论中，如果满足局域相位对称性就必须满足局域电荷守恒，这意味着必须引入光子以及与这些光子相互作用的特定形式。我们没有其他选择，光子必须存在，而且光子的出现是很自然的事情。理论中的相互作用是由它必须遵从局域相位对称性（即规范对称性）决定的，这种思想被称作规范原理。

物理学家们发现，对称性与相互作用之间存在着极其紧密的联系。量子场论提供了基本结构，却没有明确定义量子场的相位。如果允许在不同时空点上的相位有不同的值，并且要求主方程必须满足相位变化的对称性，相互作用就会很自然地出现。因此可以以不含相互作用的量子场论为起点，加上相位对称性，并且规定它符合规范原理，结果就会得到包含相互作用的量子场论——即所谓的规范理论。相互作用是由一种新引入的场（一种规范场）来传递的。在 QED 理论中，规范场就是光子。除了 QED 理论之外，电弱相互作用理论和 QCD 理论的情况也完全相同。不可思议的是，规范不变性以及一些从规范不变性中衍生出来的特殊性质是满足重正化要求的。如果不是这样，规范理论就不可能被看做描述物理过程的基本理论。

差点忘了说一句，尽管人们在很早以前就已经了解规范理论的各个部分，但 QED 理论并不是按照以上的思路建立的。QED 理论中的对称性太简单了，只有在遇到了更复杂的问题之后，理论物理学家们才真正了解到 QED 理论数学结构的重大意义，QED 理论仅仅是一个最简单的情况。

那么 QED 理论中的对称性到底是什么？作为一个仅包含自由量子电子-正电子场的理论，只要规定它满足最简单的可能局域相位变化对称性，就可以把 QED 理论构建出来。“最简单的可

能”是指：相位是一个普通的角度，范围为 0°～360°；一个相位和另一个相位结合的方式，同平面圆周上的转动结合方式相同。后者形成了一个简单平面上转动的 SO（2）群，而相位因子形成了一个数学上等价的群——U（1）群，它是一种内部抽象空间内的转动群。与 QED 理论相关的群就是转动群 U（1）群。

这种群只有一个生成元，它可以通过改变它唯一的参数——旋转角——而变化。对于一种全局 U（1）对称性，它唯一的生成元等价于只有一种荷守恒，即电荷守恒。如果规定 U（1）对称性满足局域对称，并且应用规范原理，那么在 QED 理论唯一规范场的变换性质中就会出现“旋转角”，即光子。

有一点十分重要，虽然电子与光子存在于四维时空中，并在四维时空中运动，但 U（1）对称性对应的“转动”并不是一种在四维时空内的真实转动。U（1）群的转动是一种与场相联系的抽象空间中的内部转动。之所以称之为转动，是因为它的数学表述与转动的数学表述相同，而不是因为电子上真的存在一个小球，在真实空间中做着物理上的转动；它与电子所受的电磁力有一些关系。

光子之间没有相互作用——在物理上讲，这是因为它们不带电荷；而从数学上讲，是由于 U（1）群是一种阿贝尔群。这种说法眼下看起来并无新意，但是对于 QCD 理论来讲就完全不同了——它所对应的群并不是阿贝尔群，因此顶替了光子位置的那些家伙相互之间存在相互作用。

对称性与守恒定律

规范理论的很多部分是经典的，而非量子的。像改变电磁势这种规范变换就是很明显的例子。全局对称性与电荷守恒之间的关系在经典电磁学中也可以看出来。

一个很强的、在经典物理中也适用的定理，揭示了这种联

系。这一定理，在其他场合，主要揭示一些我们不太熟悉的“荷”与相应对称性之间的关系。这个定理被称作诺特定理（Noether's Theorem），是根据德国哥廷根大学的数学家艾美·诺特（Emmy Noether）的名字命名的，她在1918年提出了这个定理。诺特定理以假定自然界遵从最小作用量原理为前提。它还用到了高斯定理（Guass' theorem），高斯定理以卡尔·弗里德里希·高斯（Carl Friedrich Gauss）的名字命名，用以阐明一个区域中某种东西产生或消失与这种东西在环绕这个区域的边界上的总流量之间的关系。接下来可以通过引进一种对称群，得出著名的守恒标量——电荷，电荷随时间的变化率，与流出或流向电荷所在区域的电流强度相一致。电荷对应着U（1）对称，与之相应的流就是电流。如果引入其他对称群，那么“荷”与“流”就应该是其他种类的荷与流，它们同电荷与电流没有任何关系。

正是诺特定理清楚地说明了空间变化下的对称性是怎样等价于动量守恒的，以及时间变换下的对称性是怎样等价于能量守恒的。在其他一些我们不怎么熟悉的例子中，也有类似的情况。

在量子场论中，规范对称性所呈现出的意义，要比在简单量子理论中出现的任何情形下都重要得多。根据诺特定理，全局相位不变性会像预期的那样得到一种守恒的荷。但在量子场论中，这种“荷”很明显是指粒子的数目减去反粒子的数目；在QED理论中，就是电子数减去正电子数。正是这种粒子数守恒，在QED理论中表现为电荷守恒。但量子场论并没有说明电荷为什么是量子化的。某个电荷总是由许多基本电荷单元组成，到目前为止，唯一能够解释为什么会出现这种情况的只有“大统一理论”。大统一理论是一种将电动力学、电弱相互作用以及强相互作用统一起来的理论框架，现在还缺乏足够的实验支持。

将规范变换这种数学工具应用到经典方程中具有重大意义，

因为放到路径积分理论中的主方程都是经典的。因此很明显，就像经典理论一样，所得到的量子理论也清楚地依赖于对称性。严格来讲，“规范理论”也能够得到非量子的结论，但在实际中，规范理论都是满足局域相位变化对称性的量子理论。

对于 QED 理论，相位对称性就是只有一种转动情况下的对称性，量子场就是电子-正电子场。在其他理论中，存在更加奇异的对称性，这就意味着存在更加奇异的对称群和更加奇异的转动，并且也会有不同的量子场。每一种转动角对应着一类规范场，在 QED 理论中这种规范场就是光子；并且对于每一种转动都会存在与之对应的生成元，生成元反过来又决定了一种守恒的荷，在 QED 理论中这种荷就是电荷。

由于科学家们所研究的电动力学理论、电弱相互作用理论以及强相互作用理论都是规范理论，因此它们三者之间存在许多共同特征。最简单的一点就是：至少它们的规范场都应该是没有质量的；至少大家都知道光子是没有质量的，因为一旦这三种理论中哪一个主方程含有一个有质量的规范场项，那么那种特殊的规范对称性就会遭到破坏。在所有的情况下，规范原理（强制其满足一种局域相位对称性）决定了相互作用，并且在那些具有更奇异的对称性的理论中，规范原理对它们相互作用的耦合常数施加了重要的限制。

在这三种理论中，所有的基本物质粒子都是自旋为 1/2 的费米子，就像电子和正电子都是自旋为 1/2 的费米子一样。规范理论说明了允许物质费米子之间产生相互作用的规范场为什么必须是向量场。在 QED 理论中这是很容易看出来的，因为它的规范场由一个标量的电势和一个三维向量磁势组成，所以它的规范场是一个大家熟悉的四维向量场。另一种解释向量场必然出现的方法，是用电流的观点来看待电荷碰撞。一个运动中的带电电子或正电子实际就是一种电流，而流是四维向量。发生相互作用后，

电流会通过某种方式经由电磁场发生改变，因此电磁场要完成这项工作就必须具有向量的性质。由于一个四维向量场与自旋为 1 的场具有相同的转动变换性质，因此有关弱相互作用、强相互作用以及电磁相互作用的规范理论都认为，相互作用是通过交换自旋为 1 的规范场来实现的。规范场也有可能具有更高的自旋值，描述引力的规范理论应当具有自旋为 2 的规范玻色子——引力子，引力子传递纯吸引的万有引力——两个受万有引力作用的物体永远也不会互相排斥。

在量子场论中，自旋值为 0、1、2 等整数的粒子必须遵守满足玻色-爱因斯坦统计的量子法则，因此称作玻色子。这是服从狭义相对论因果关系的必然结论：只有当光信号能够在两个事件之间传递信息时，两个事件才有可能产生依赖关系。与之类似，具有半整数自旋值的粒子也必须遵守费米-狄拉克统计，因此是费米子。这些内容在本质上就是自旋统计定理，它们的出现在很大程度上归功于泡利，他在 1940 年发表了一篇有关两者之间联系的关键论文。量子场论为自旋与统计之间的联系提供了唯一已知的解释，这是量子场论成功的标志之一。

因此，在规范理论中，携带力的粒子都是玻色子。费米子正是通过这些携带力的粒子发生相互作用。费米子就是宇宙中“其他”的组成部分，它们是“物质”。只有当费米子符合特殊的统计规律时，它们才能以独特而众所周知的性质共同组成更复杂的东西。一个不存在相互作用的宇宙将是一片黑暗，只存在着不能结合到一起的费米子，因此它们也不能组成我们周围所能看到的一切东西。

QED 理论

QED 理论是多么美妙呀！QED 理论是有史以来最优秀的理

论，如果忽略重正化和微扰论的侵扰，它堪称是一部集所有理论的优雅于一身的杰作。

用现代术语表示，QED 理论是一种具有局域阿贝尔规范对称性的量子场论。它的规范群是 U（1）群，QED 理论在局域规范变化下的不变性，导致要引入一种没有质量的规范玻色子，也就是光子；与之对应存在着一种守恒的荷，叫做电荷。这种荷存在于一种属于费米子的粒子-反粒子对之上，即电子-正电子对。QED 理论是一种关于光与物质相互作用的量子理论，它符合狭义相对论，并且能够解释两种费米子和光子的诞生和湮灭。

1927 年，狄拉克最早发表了关于 QED 理论的论文，这篇文章不仅为 QED 理论打下了基础，而且标志着量子场论的出现。在随后的几年中，量子场论被费米用到了关于弱力的理论中，汤川秀树也试图用它来理解强力。

但是在直到第二次世界大战爆发的这一段时期内，物理学家们放弃了在这一方面的努力，计算真实物理系统的尝试一直受到发散问题的困扰。在战后的几年中，发散问题很快被解决了，这些进展部分地受到了两个重要实验结果的鼓舞。

在 1947 年，哥伦比亚大学的拉比（Isidor Rabi）同两个研究生约翰·内夫（John Nafe）和爱德华·纳尔逊（Edward Nelson）一道，完成了一项有关氢光谱精细结构的非常精密的实验。他们研究了两个特殊的能级，重要的是，这两个能级间的间隔与电子和质子磁矩之间相互作用的差别有关。拉比和他的小组发现，他们测量到的能级间隔同现存理论计算出来的不符。根据耶鲁大学理论物理学家格雷戈里·布赖特（Gregory Breit）的意见，对于他们实验的结果，一种可能的解释是，电子磁矩的假定值不正确。

同样是在哥伦比亚大学，同样也是在 1947 年，威利斯·兰姆（Willis Lamb）和他的研究生罗伯特·热泽福特（Robert Re-

therford）做了另一个有关氢光谱精细结构的实验。这是人们第一次测量到兰姆位移，但是根据狄拉克的理论，氢光谱不应该具有这种性质。

这两个实验结果都在 1947 年 6 月 2～4 日举行的“沙尔特岛”会议（Shelter Island conference）上公布了。会议在纽约州沙尔特岛的 Ram's Head 酒吧举行，24 位在理论物理方向进行研究的美国顶尖物理学家汇聚一堂，使得这次会议极富传奇色彩。兰姆和拉比的结果引起了大家新的兴趣，这些精确的结果为理论物理学家提出了新的挑战，也为他们的计算结果提供了精确的参照。这次会议预示着 QED 理论的发展进入了新的阶段。

两年之后，哈佛大学的施温格和东京教育大学的朝永振一郎通过重新定义质量和电荷的方法，消除了 QED 理论中可恶的发散（至少在低次项是成立的），从而分别独立地提出了几乎完全一样的重正化方法。在 1948 年，施温格就能够计算出电子磁矩的场论修正值，他得到的结果同新的实验结果相吻合；同一年，就在比施温格稍早一点的时候，朝永振一郎和他的同事第一次公布了用场论计算兰姆位移所得到的正确值。在康奈尔大学，费曼用一套基于路径积分的、与他们完全不同的方法，也得到了相同的结果，他用截止限制了发散。他的计算结果发表在 1947～1950 年的一系列论文当中。1949 年，戴森证明了费曼的方法实际上与施温格和朝永振一郎用的方法完全等价，他在同一年还证明了重正化方法对所有次项都成立。费曼、施温格和朝永振一郎因他们在这方面的工作分享了 1965 年的诺贝尔物理学奖。

兰姆位移现象的产生，一部分是由于电子自能（或真空极化）导致库仑定律发生改变，见图 3.8 的图（b）。原子中的电子通过交换光子而被束缚在某个区域：根据 QED 理论，光子能够在短时间内自发地产生电子-正电子对，也就是发生真空极化，这起到了屏蔽电子和原子核之间相互作用的效果；正是这种屏蔽

效应改变了库仑定律。更高次项的费曼图对应着额外的屏蔽效应。简单来讲，兰姆位移起源于两个特殊的电子态，它们本应该具有相同的能量，但实际情况却不是这样。QED 理论的修正显示，每一个态的能量都有微小的偏移，这部分上是由真空极化造成的。但两个态中每一个态的能量偏移是不一样的，因此它们不再具有相同的能量。按照 QED 理论，原来的单独存在的一条能级变成了两条分裂的能级。经 QED 理论修正过的结果是正确的，兰姆和热泽福特在实验中测量到的，正是这种由能级偏移造成的微小能隙。兰姆被授予了 1955 年诺贝尔物理学奖，来自哥伦比亚的物理学家库什（Polykarp Kusch）和他分享了这一奖项。

另一方面，电子的反常磁矩对应于顶点相互作用图，图 3.8 的图（c）表示的就是最简单的顶点图。狄拉克理论所预言的磁矩值适用于没有内部结构的点状电子。而反常磁矩——上面提到的磁矩值和实验测量值之间的差别，说明了电子的质量和电荷确实存在某种分布。QED 理论用电子周围围绕着虚光子云来解释这种分布，这些光子不断地被电子发射又被重新吸收。在任一时刻，电子把一部分质量-能量借给了虚光子，在效果上改变了电子质量相对于电荷的分布，这样就造成了磁矩的变化。可以肯定的是，图 3.8 的图（c）所示的费曼图，即电子磁矩的一级辐射修正，提供了一个值为 $\alpha/2\pi$ 的因子。这个因子的数值是 0.0011614，可以将它同实验值作对比，根据在本章开始所引用的数据，实验上测得的电子反常磁矩的值是 0.0011596。施温格在 1948 年计算出来的就是因子 $\alpha/2\pi$ 这一部分。更高次项的修正，包括虚光子反过来又生成的虚电子-正电子对的贡献，不断刷新着计算结果。实验与理论在 4 个圈的精度上已经十分吻合了，这个事实不但给 QED 理论提供了已知的最好验证，也极大地支持了重正化方法。

库什同亨利·弗利（Henry Foley）在 1948 年一道测量了电

子的磁矩，并且正如布赖特猜想的那样，他们确认了电子磁矩的值确实同狄拉克理论预言的不同。电子的反常磁矩支持了拉比的光谱测量结果。

兰姆位移和电子的反常磁矩（用场论的语言表述）都是因为电子自发发射以及重新吸收光子，或者由光子产生或吸收电子-正电子对，从而造成物理性质改变的结果。如果从粒子发射的角度说，这种现象就被称作“辐射修正”，它的意思是指由于发射粒子而改变了物理过程，否则电子仅仅是一个电子，而光子也仅仅是一个光子。

许多过程其实就是辐射修正过程，场论预言了许多奇特的过程，这些过程在比较简单的量子理论中不可能发生，因为它缺乏描述粒子产生和湮灭的能力。

光子之间的碰撞和散射就是一种这样的过程（稍后在适当的时候我们会再次提到它）。乍一看上去，这是不可能的事情，因为两个光子都不带电荷，那么它们之间怎样发生相互作用呢？答案是每个光子都变为电子-正电子对，图 3.8 的图（e）所示的光子-光子散射图描述了这个过程。这个由两个光子碰撞又得到两个光子的过程，表示为 $\gamma+\gamma\rightarrow\gamma+\gamma$，这种现象人们已经观测到很多次了。

例如在 20 世纪 80 年代，德国北部 DESY（the Deutsches Elektronen-Synchrotron in Hamburg，德国汉堡电子加速器）的实验者们用 DORIS（正负电子双存循环）对撞机中的“水晶球”探测器研究了这个过程。电子和正电子通过发射光子产生相互作用——在适当的条件下，电子和正电子都能够产生真实光子，因此实际研究的碰撞过程是光子与光子的碰撞过程。“水晶球”是一种碘化钠晶体球，它覆盖了几乎整个发生相互作用的区域，它能够作为探测器依赖于一种辐射过程。当光子进入晶体时，就会产生电子-正电子对。作为一种中间过渡物质，它们会通过一种

称作韧致辐射（bremsstrahlung，或者 braking radiation）的过程再次生成光子。这些光子反过来又会通过电子对产生过程，产生更多的电子-正电子对。这些过程反复进行，就会产生一种电磁雪崩，导致碘化钠放出可以测量到的闪光，发出这种闪光就表示原始的光子已经到达探测器。

一种奇特的 QED 辐射过程为观测反原子奠定了基础，第一个观测到的反原子事例发表于 1996 年。在德国苏黎世研究中心核物理研究院工作的沃尔特·奥勒特（Walter Oelert）和他的同事一道，利用 CERN 的低能反质子环（LEAR，the Low Energy Antiproton Ring）生成了数目可观的反氢原子。我们都知道氢原子由一个质子和一个电子组成，而反氢原子大致也是通过完全一样的方式，由一个反质子和一个正电子组成。这只是一种假定，因为人们制造反氢原子的目的之一就是想要看一下，反氢原子是否真的同普通的氢原子符合同样的物理规律。

早在两年之前，SLAC 的斯坦·布鲁斯基（Stan Brodsky）和查尔斯·芒格（Charles Munger）以及位于智利瓦尔帕莱索（Valparaiso）的费德里科·圣玛丽亚大学的伊凡·施密特（Ivan Schmidt）就一起设计了一种生成反氢原子的方法。奥勒特和他的小组在 LEAR 上将他们的计划付诸现实，他们让反质子通过一束氙气，之所以选择氙气是因为它是一种具有很重原子核的气体。在氙原子核附近，一个反质子会产生一个光子，光子又会生成一个电子-正电子对。偶尔，会有正电子被原来就存在的反质子俘获，这样就生成了反氢原子，如图 3.15 所示。

所有的事例都说明，QED 理论为 QCD 理论的发展做了大量的准备工作，因为人们认为，QED 理论中的许多过程同 QCD 理论中的过程类似。例如在 QCD 理论中，理论预言一个 π^+ 和一个靶原子核之间的碰撞会导致一种叫做 D^+ 介子的粒子的产生。因此在 QCD 理论中也存在与 QED 理论中相类似的过程，但是后者

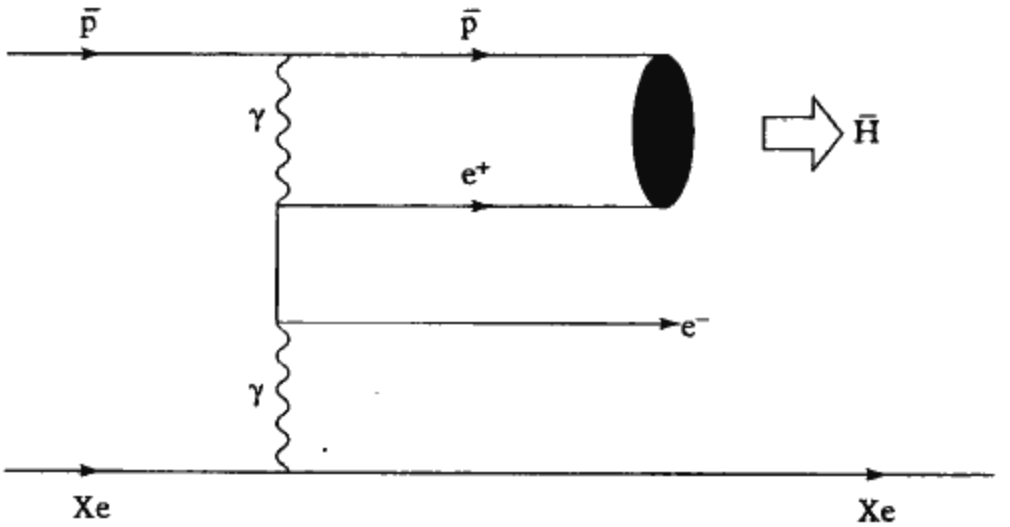


图 3.15 通过反质子 p̄ 同氙原子核 Xe 发生量子电动相互作用来产生反氢原子 H̄ 的过程

可以通过实验检验得更精确，也更容易理解。

此外，QED 理论肯定了场论形式的价值，证明了重正化方法的强大威力，表明了能从微扰论中得到有意义的结论，并且还展示了规范原理的作用，所有这些都可以直接应用到 QCD 理论之中。

对于 QED 理论来讲，电弱相互作用理论除了建立起一套关于弱力的规范场理论之外，还进行了其他一些工作——它是一套既描述弱相互作用又描述电磁相互作用的理论。QED 理论是基于U（1）群的；电弱相互作用是基于两个群的乘积，写作 SU（2）×U（1），其中 U（1）部分就是 QED 理论中的群，SU（2）部分来自于弱相互作用。电弱理论只有在弱力和电磁力统一的情况下才起作用。独立的 QED 理论仍然有效，它是一套更普遍的理论在低能情况的表现。但是没有独立的弱相互作用理论与之相对应。

将弱力也统一进来的电弱理论是基于 SU（2）×U（1）群

的，描述强力的 QCD 理论则基于 SU（3）群，它们在粒子物理学家的头脑中的优越地位毋庸置疑，它们共同负责解释自然界中四种基本相互作用中的三种。物理学家们好像还要让这种被广泛接受的思想在大家脑海中的印象更加深刻，将这些理论称为“标准模型”。理论物理学家们都有一种直觉，他们认为完成统一大业的下一步，就是要将电弱理论和 QCD 理论融合到一起。到目前为止，这项计划还没有圆满完成。

既然到目前为止，QED 理论已经完成了说明规范场论怎样运作的任务，那么现在是继续我们旅程的时候了。QCD 理论中的对称性从何而来？这场游戏中，谁又是最关键的角色呢？

第四章 向 QCD 进军

强　力

如果中世纪的炼金术士们首先发现了核物理学，他们会做得更好。元素的化学性质由原子中的电子数和电子排列所决定，原子中的电子数一定要等于原子核中的质子数。因此要把一种元素转变为另一种元素，就意味着要改变质子数。无论通过什么化学手段都不能将铅变为金，但早期的炼金术士们只了解化学。他们还应该了解一些别的知识。

世界上第一个成功的炼金士，是伟大的新西兰实验物理学家欧内斯特·卢瑟福（Ernest Rutherford，图 4.1）。从炼金士的标准来看，他开始的目标很保守，只是想将氮变成氧。在英国曼彻斯特的一次实验中，他利用放射性镭产生的天然 α 粒子轰击氮，完成了这一壮举，实验结果于 1919 年公布。这是人类有史以来第一次实现用人工改变元素，也是第一次人工核反应。有一点讽刺的是，这明明是一次物理实验，但这位“核炼金士”卢瑟福先生那时已经于 1908 年因“在分解元素以及放射性物质的化学性质方面的研究”获得了诺贝尔化学奖。

卢瑟福确定，反应中产生的长距离辐射粒子是氢原子核。到1920年，他已经开始把它们称作质子，并且他还了解到，质子是原子核的一种组分，卢瑟福将其称为“自然的结论”。

卢瑟福用α粒子轰击氮所得到的结果，只是他研究计划的一部分，他打算用射线处理一系列较轻的元素。目的就是让α粒子与原子核之间的距离离得尽可能近，以便研究在原子核周围，静电学中的库仑定律是否成立；或者反过来，看看原子核周围是否存在电子。库仑定律表述了两个带电物体之间的经典静电力；如果两个电荷之间的距离变成原来的2倍，它们之间的相互作用力就会减弱为原来的1/4。这就是被物理学家们称作平方反比定律的一个实例。

由于比较轻的元素，其原子核所带的电荷也较少，因此射入的α粒子能够离原子核更近。从这种观点来看，最轻的元素当然也就是最好的，这种元素就是氢。但是当卢瑟福用α粒子轰击氢原子核时，他发现α粒子的散射规律不符合当时的简单理论。简单理论认为α粒子和原子核都是点电荷，它们通过平方反比规则相互作用。他观测到，α粒子和原子核之间相互作用的强度和方向，在相距很短时会发生奇怪的变化。至关重要的是，卢瑟福选择了用静电学来处理他的结果，因此他把这种效应归结为α粒子内部具有复杂的结构，并且存在于原子核内部的真实电子发生了变形——以今天观点来看，电子是一种点粒子，卢瑟福的这种想法是十分奇怪的。

在调到剑桥之后，卢瑟福让詹姆士·查德威克（James Chadwick，图4.2）和麦吉尔大学的艾蒂安·比勒尔（Etienne Bieler，图4.3）对氢散射α粒子的问题进行了详细研究。查德威克和比勒尔不断寻找有关α粒子大小和形状以及α粒子周围力场的详细信息。同卢瑟福所做的实验一样，他们也用从放射源射出的α粒子轰击氢靶，但是他们使用了一种更高级的显微镜，来给硫化锌

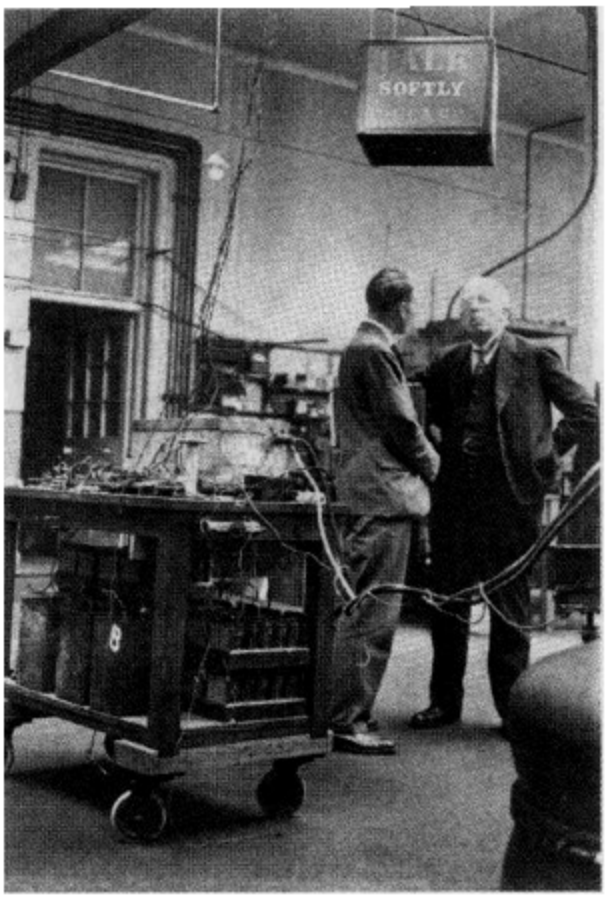


图 4.1 轻声交谈。这张照片拍摄于 1932 年的剑桥大学卡文迪许实验室。生于新西兰的欧内斯特·卢瑟福（面对照相机者）是第一个意识到原子含有带正电荷的原子核的人。他还首次实现了人工将一种元素转化为另一种元素的过程，他因在放射性方面的工作获得了 1908 年的诺贝尔化学奖。卢瑟福认识到，质子是原子核的一部分，并预言存在中子。图中“轻声交谈”的标志是为了避免声音干扰周围精密仪器的灵敏度而设，因为卢瑟福的大嗓门十分出名。图片来源：剑桥大学卡文迪许实验室

屏上产生的“氢粒子”发出的亮光计数。他们对氢粒子的产生怎样随 α 粒子的速度变化、氢粒子的角分布以及散射角固定时氢粒子的强度等问题都做了系统的研究。

他们的研究结果发表在 1921 年的《哲学杂志》（*Philosophical*

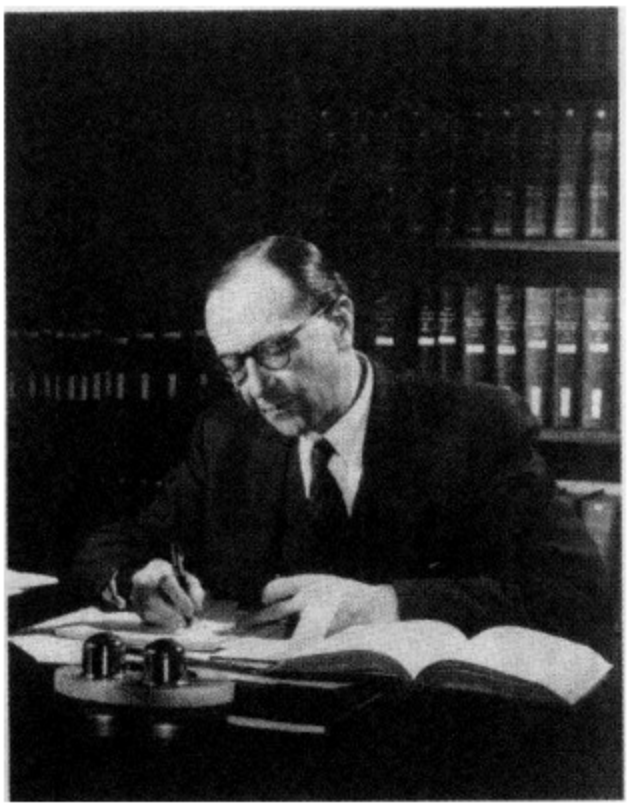


图 4.2　查德威克利用氢对 α 粒子的散射过程揭示强相互作用的存在。1913 年，查德威克搬到了柏林，同汉斯·盖革一同工作，但是随着第一次世界大战的爆发，他在一座赛马场的马厩中被关押了 4 年。早在 1914 年，查德威克就在 β 衰变方面作出了重要的贡献。他在 1932 年又发现了中子，因此获得了 1935 年的诺贝尔物理学奖。图片来源：剑桥大学卡文迪许实验室

Magazine）上，实验清楚地说明，尽管静电学中简单的平方反比定律对低能 α 粒子仍近似成立，但是它对高能 α 粒子不成立。他们提出了几种可能的 α 粒子模型，最终确定 α 粒子最像一个长短半轴分别是 8×10^{-15} 米和 4×10^{-15} 米的弹性“扁球体”。他们在报告中写道：

“根据这种观点，一个射向 α 粒子的氢粒子，会在普通

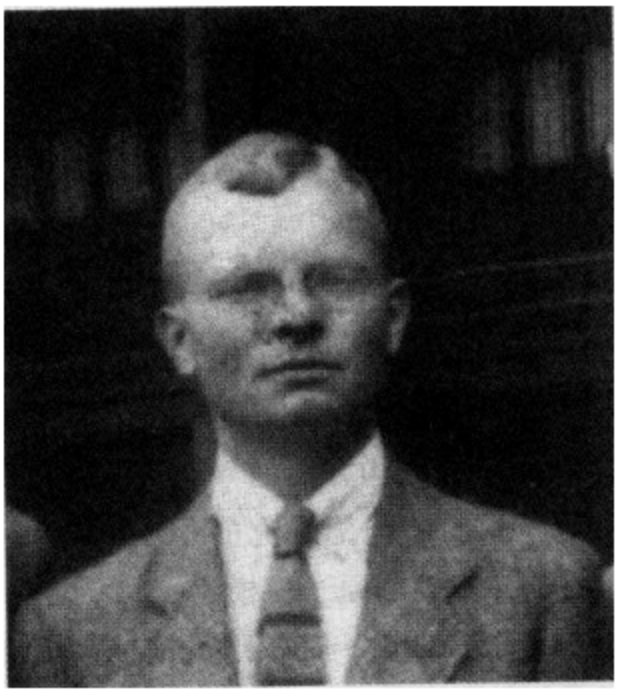


图4.3 1921年的艾蒂安·比勒尔，拍摄于剑桥大学卡文迪许实验室。他和查德威克一起首次观测到了强力。图片来源：剑桥大学卡文迪许实验室

静电学中的平方反比定律下运动，直到它到达一个半径如上所示的球面。这时它会遇到一种极强的力场，就好像从一个很硬的弹性物体上反弹回来一样。”

他们已经窥探到了强力的迹象。查德威克和比勒尔的工作被认为是发现强力——也就是让原子核的各种成分聚在一起的那种黏合剂——的最早证据。

质子与中子现身

早在20世纪20年代，卢瑟福就推测存在一种具有一个单位质量但不带电荷的“原子”，卢瑟福将其设想为一种由电子和氢原子核组成的“中性粒子对”。经过12年苦苦搜寻，查德威克通

过用α粒子轰击铍发现了这种东西，他适时地把它称作“中子”。在早期，人们只知道中子的质量大约与质子相同，它所带的电荷数为零，与氢或者氮之类的原子碰撞的时候会引起这些原子反冲。有人认为中子是一种质子-电子态，也有人认为它本身应该是一种基本粒子。

尽管很难用实验直接证明，但在中子被最终发现以前，人们已经普遍认为原子核由中子和质子组成。电子-质子模型同狭义相对论和量子力学明显矛盾，正是这些困难激发了海森伯（Heisenberg）在原子核结构方面的兴趣。海森伯抓住了进一步理解原子核的关键，那就是查德威克的中子，他在1932年提出了一套以中子和质子为基础的原子核模型。

海森伯的论文分三部分，他在原子核结构中应用了简单的量子力学，这标志着现代原子核结构理论的开端。尽管他仍保留了中子由电子-质子对组成，而不是一个独立粒子的思想——这从现代物理的观点来看是一个瑕疵，但后来的事实证明，这确实是一个对后来发展具有关键意义的思想。

1932年初，哈罗德·尤里（Harold Urey）和他的同事发现了氢的一种同位素——氘，这对海森伯的工作具有重大意义。按照现在的说法，一种元素的所有同位素，其原子核中都具有相同的质子数（因此原子也具有相同的电子数和化学性质），但具有不同数量的中子数。氘包括一个质子和一个核外电子（这就是“普通”氢原子的成分），但除此之外，它的原子核中还含有一个中子。顺便说一下，与氘不同的是，大多数同位素都没有自己的专门名字。

对于海森伯来讲，中子在某种程度上是一个质子-电子对。因此他认为，如果一个中子和一个质子结合，那么中子中的电子将会被两个质子所共享。这样，电子就能通过“交换力”把两个质子结合到一起，就像几个原子共同结合成一个分子一样。通过这种方

式，中子和质子可以结合在一起形成氘原子核。数量更多的质子和中子也可以以同样的方式黏合到一起，形成更重的原子核。他根本没有考虑中子同质子间相互作用力的精确形式是什么样子，他的这一猜想导致了一个大行当，随后很多人尝试了各种各样的数学形式，来表示这一相互作用力。

从后来的发展来看，海森伯工作中最重要的一点，就是他认为中子和质子在本质上是一种东西；他提出的公式中，包含了一种数学表达式，可以用来描述质子和中子间的相互转变。他提出了一个新的变量，这个变量可以取两个值：一个值对应着质子，另一个值对应着中子。用现在的话说，海森伯引入的是这样一个概念：质子和中子是某一种更基本的物质的两个态。若干年后，人们将这种物质命名为核子。如今，核子这个词实际上是一个既表示质子又表示中子的统称。

海森伯模型是用来解决质子与中子的结合问题的。在诸如两个中子或两个质子怎样结合到一起的问题上，该模型不甚明了。实际上，一开始人们还不敢肯定两个中子或两个质子是否能够结合。但是到 1936 年，物理学家们已经意识到，两个中子或两个质子不仅能够结合，并且如果忽略电磁力的话，它们之间的结合强度同一个质子与一个中子结合的强度是一样的。简而言之，将核子结合到一起的力与电荷没有任何关系。

类比一下的话，会发现核子有一点儿像人：一个人是一种单独的客体，会处于两种可能的状态之一——比如说男或女。许多国家的法律都宣称，男女是应该受到平等对待的；在这种情况下，法律只对人起作用，男人和女人之间没有任何差别——这里存在一种对称性。但是从生物学的角度或者其他方面来看，一些人可能会说男人和女人是不一样的，因此相应地行为也会不同。这就像核子的两个态，它们在电磁学问题上会产生不同的行为。

电荷与原子核的结合无关，这可以用海森伯在 1932 年引进

的新方法和数学量自然地表现出来。这个新的量被命名为同位素旋（isotopic spin）；1937 年，把群论引入量子物理的尤金·魏格纳（Eugene Wigner）将其简写为同位旋（isospin）。被海森伯用来描述质子-中子互相转化的数学方法就是群论，将量子理论和群论应用到原子核上，就可以为世界带来一种全新的物质性质——同位旋。几年之后，事实证明了确实存在同位旋对称性，它对于理解强相互作用以及将基本粒子进行归类，都是至关重要的。

海森伯的质子-中子模型推翻了统治多年的原子核质子-电子模型。但它还是无法解决科学家们在理解原子核内部状况时所面临的所有困难。困难之一就是怎样理解 β 衰变（beta decay）——这是一种能发射电子和反中微子的自发辐射衰变现象。

弱相互作用小插曲

恩里克·费米（Enrico Fermi）在 1933 年用一种引人入胜的方法解决了这个问题。费米的理论引入了两个关键思想，这不仅对理解 β 衰变十分重要，而且对原子核与粒子物理的整体发展都具有重大意义。这两种思想分别是：衰变是一种全新的力造成的；像电子和中微子这样的粒子能够产生出来，而不是早已经存在，仅仅从一个现有的房子里被赶出来。费米引入的新相互作用力，现在被称为弱力，他提出的理论是现代电弱相互作用理论的前身。

费米考虑到，如果将他的理论变为一种量子场论，就可以让电子这样的粒子产生出来，电子产生的方式同电动力学中光子产生的方式类似。这种描述电子产生的力学不需要将电子预先藏在原子核或中子里——那时，人们还都认为 β 衰变中的电子来源于中子。

粒子产生的思想要求存在一种叫做中微子的粒子。在研究 β

衰变的过程中，有一个谜团长期存在，那就是衰变产生的电子所具有的能量有一个连续的范围，而不是某一个固定的值。这种现象不符合能量守恒定律，玻尔等一些人认为这是个很严重的问题。1930年，泡利提出了一种同样基本的思想，他认为衰变过程在放出电子的同时，也同时放出了一种看不见的新粒子，这种新粒子也会带走一些能量，因此整个过程就够满足能量守恒了。这是一个艰难的选择：是在原子核内部放弃能量守恒，还是在保留能量守恒的前提下，引入一种神秘的新粒子呢？

费米选择了后者。确实，只有当一种相关粒子伴随电子产生时，他的理论才符合实际情况。用现在的话说，费米的理论将原子核的β衰变现象解释为中子通过弱相互作用衰变为一个质子、一个电子和一个反中微子。电子和反中微子并没有呆在原子核中等着逃出来，而是在衰变过程中从无到有产生出来的。

（反）中微子确实很难观测到，这就是为什么虽然人们在研究衰变方面投入了巨大的实验力量，但到费米提出理论的时候，还没有人能观测到它。弗雷德里克·瑞恩斯（Frederick Reines）和克莱德·科万（Clyde Cowan）将核反应堆作为中子源，最终在1956年发现了中微子。瑞恩由于在这方面的成就，和别人一起分享了1995年的诺贝尔物理学奖。

尽管费米的理论有许多成功之处，并且为物理学的发展带来了重大影响，但最终还是失败了。在低能状态，例如只有几百兆电子伏的情况下，费米的理论与实验符合得很好。但是在更高的能量下，计算结果会出现明显的问题，会得出大得不可思议的散射截面值。问题在于费米的理论违反了一个关键原理，即幺正性原理——将系统处于所有末态的可能性累加起来，得到的总概率值应该为1。并且费米的理论还有另一与幺正性紧密相关的问题，即它不能重正化。在电弱理论中，这两个问题都得到了适当的解决。

在人们将费米理论作为一套解释核力的理论，并认为它可以解释中子与质子的结合的时候，它还没有遇到那么多挑战。但是不久，人们就清楚地发现，费米的理论不可能完成第二项任务，因为它所描述的相互作用力强度没有那么大。但就在海森伯等一批物理学家努力解决这一问题的时候，一条关键的线索已经隐藏在文献之中。

这条线索是由汤川秀树在 1935 年提出来的，他提出在原子核中将质子与中子结合到一起的相互作用力，应当由一些比较重的交换粒子传递（图 4.4）。这种思想对于当时还处于萌芽状态的交换力思想是一项全新的贡献，它最终在人类理解基本相互作用的问题上引发了巨大的变革。汤川秀树提出的携带相互作用力的重粒子——π 子，于 1947 年在实验中首次亮相。汤川秀树工作的另一个特征，就是承认中子与质子之间的相互作用力是一种比电磁相互作用力强许多倍的全新作用力。强力于 1921 年首次在实验中被观测到，在汤川秀树的努力下，它最终在基本作用力的框架中占有了一席之地。

加速器的曙光

在 20 世纪 30 年代，人们不仅开始见识到弱力、强力以及其他许多关键的理论思想，还见证了粒子加速器的诞生。加速器的产生实际开始于 1932 年，这一年剑桥大学的两位物理学家约翰·考克饶夫（John Cockcroft）和欧内斯特·沃尔顿（Ernest Walton）赢得了分裂原子的比赛。他们的粒子加速器通过一系列电容器和整流器，将质子加速到 700 千电子伏。他们用制成的快质子束轰击锂，最终得到了 α 粒子对，世界上第一次依靠人工加速粒子完成的核反应实现了。

许多其他的物理学家也在建造加速带电粒子的仪器。在普林斯顿大学，罗伯特·范·德格拉夫（Robert van de Graaff）研制

图 4.4 朝永振一郎（左）、汤川秀树（中）和坂田昌一（右）。朝永振一郎因他在 QCD 理论方面的工作获得诺贝尔物理学奖，汤川秀树因提出强相互作用的介子交换理论而获诺贝尔物理学奖。坂田昌一提出了一种以质子、中子和 Λ 粒子为基本砖块的重子和介子模型。图片来源：筑波大学朝永纪念室，美国物理学会艾米利奥·塞格利可视资料馆（AIP Emilio Segrè Visual Archives）

出了静电发生器。静电发生器现在还是以他的名字命名的。在 1931 年，范·德格拉夫最初所用的机器能产生约 1.5 兆电子伏的静电压，他的机器像考克饶夫和沃尔顿的电压倍增器一样，通过电压降的简单原理来加速粒子。

而在加利福尼亚，物理学家们使用了不同的加速方法。在加利福尼亚大学，与考克饶夫和沃尔顿同时代的欧内斯特·劳伦斯（Ernest Lawrence）和他的学生斯坦利·利文斯通（Stanley Livingston），一同研制了一种全新的机器——回旋加速器，来加速带电粒子。考克饶夫和沃尔顿并没有利用一个很大的电压降来加速电子，他们的加速器是基于当时新提出的共振加速理论建造的，粒子可以多次通过相同的较小电压进行加速。

挪威人鲁夫·怀特罗（Rolf Wideröe）在德国的亚琛建造过

一套仪器，这是人类建造共振加速器的第一次尝试。怀特罗的加速器由排成一条线的三个圆柱形电极组成，每两个电极之间都有间隔。他在中间的电极和两端的电极之间施加了一个交流电压。在某个周期的开始，中间的电极相对于另外两个电极是负极，因此一个带正电的粒子会从第一个电极向中间的电极加速。当它经过中间的电极时改变电压，这时中间的电极相对于最后一个电极是正极，因此当电子离开中间的电极后会再次加速，继续向第三个也就是最后一个电极加速前进。其中的技巧就是调整电压改变的频率，使其与电子通过中间电极的时间一致。怀特罗那台简陋的仪器，是现在所有现代粒子加速器的鼻祖。

劳伦斯偶然读到怀特罗的论文，他意识到可以用磁场使带电粒子的路径弯曲，这样粒子就有可能重新回到第一个电极，然后就可以一次接一次地循环，进行周期性的加速。这种加速器就是回旋加速器（cyclotron）。

在回旋加速器中，磁场引导带电粒子以螺旋形的运动轨迹，进入两片中空的 D 形（半圆形）电极，两片电极的直边相对放置，中间有一条很窄的缝隙。在两片电极上施加周期变化的交流电压，产生用来加速粒子的电场。调整磁场，让粒子在一个 D 形电极中沿半圆形轨道运动的时间，与电压在两个电极间反转的时间相等。这说明每次粒子到达电极之间的间隙时，都能够被再次加速。第一台可用的回旋加速器于 1931 年 1 月展示在世人面前，它直径只有 4 英寸（1 英寸约为 2.54 厘米），用一个手掌完全可以将它盖住。稍后的一台回旋加速器于 1932 年建成，它比第一台稍大一点，直径 11 英寸，当粒子向这台机器的外圈回旋时，每次会在 300 伏电压下再次加速，最终粒子的能量能达到 1.2 兆电子伏。劳伦斯、利文斯通和米尔顿·怀特（Milton White）用这台仪器重复了考克饶夫和沃尔顿曾做过的实验，他们用质子轰击锂，也得到了 α 粒子。

回旋加速器的尺寸不断增加，它很快成为核物理研究中的主要力量。但是，回旋加速器有一个内在的限制。根据狭义相对论，随着能量的增加，被加速粒子的质量也不断增加，对于能量达到25兆电子伏或者更高的质子，这种质量增加变得十分显著。这对于回旋加速器是一场灾难，因为它破坏了仪器工作所依赖的共振条件——随着粒子能量的增加，它们通过D形电极的时间会变得特别长，到达电极间隙的时间也会特别晚，因此会错过加速电场。仪器中粒子的旋转频率不再同加速电压的频率保持一致，因此粒子就不能被继续加速。对于简单的质子回旋加速器，能量的实际上限大约是25兆电子伏。

第二次世界大战结束后不久，人们提出了一种现代加速器设计的基本原理，解决该问题的方法也随之诞生了。这种思想是由弗拉德·维克斯勒（Vladimir Veksler）和爱德华·麦克米伦（Edward McMillan）分别在1944年和1945年各自独立提出来的，它被称作相位稳定原理。随着加速器中粒子质量的不断增加，可以通过改变加速电压的频率和束缚磁场，使粒子在仪器中运动的频率始终与加速电压的频率保持同步。具有适当能量的粒子会一直与加速电压同时出现在电极间隙中。那些到得太早的粒子被加速得更快，因而会沿着一条比较长的轨道运动，所以下一次到达间隙的时间会晚一些。而与之相反，那些原来到得比较迟的粒子则得到较少的加速度，会有一个比较短的运动轨道，下一次就会提早一些到达。通过这种方法，被加速的粒子就会变得步调一致。这种被加速的粒子和对粒子进行加速的电压之间相对时间保持不变的性质，被称作相位稳定性。

一台回旋加速器要想使粒子达到更高的能量，一种方法就是保持间隙宽度和磁场强度固定不变，而逐渐减小施加在间隙上的加速电压的频率。结果就得到了一种新的仪器，它有一个新名字，叫做同步回旋加速器（synchrocyclotron）。

但随着能量的增加，还是会出现不利的影响。从加速器里输出的粒子不再是连续喷涌的粒子束，而是一团一团地喷出，每经过的一个完整的加速周期就会输出一团粒子。这暴露了一个相当严重的问题，因为在回旋加速器中大部分时间内都没有粒子束，实验能够利用的粒子数量会特别少，大约是回旋加速器中粒子数的1%。

世界上第一台同步回旋加速器于1946年在伯克利投入使用，它能够将质子加速到350兆电子伏。之后人们又陆续建造了许多同步回旋加速器，其中包括位于莫斯科附近杜布纳市的原子核联合研究所（JINR，Joint Institute for Nuclear Research）的那台巨大的仪器，它能使粒子的能量达到680兆电子伏。这台机器于1954年开始投入使用，如今已经经过改进，并且被重新命名为同相加速器（Phasotron），现在仍被用于核物理方面的研究。1958年，一台600兆电子伏的加速器在CERN投入使用。尽管已经经过了多次改造，但像JINR那台加速器一样，它现在仍在运行。

但这些较大的加速器代表了这种加速方案的能量上限。在更高的能量下，粒子束就会在主路径附近以一种无法控制的方式来回摇晃。不仅如此，用来引导粒子进入螺旋轨道的磁铁也需要做得特别大，在现实中很难实现。

幸运的是，同步回旋加速器并不是实现相位稳定的唯一方法。还存在另外一种改进方式，即提高加速周期的频率，同时增大磁场。除了具有相位稳定性之外，这种方法还有另一个巨大的优点——用这种方法加速粒子的路径可以被控制在一个圆周上，而不再沿着螺旋线向外扩散。这样就只需要在一个很小的圆环状区域内施加束缚磁场，因此可以节省大量耗费在磁铁上的成本。实际上，因为只需要一个环状的区域，整个加速器可以被建成一个由环绕着引导磁线圈的管子组成的封闭圆环，在沿着管子的一个或多个点上施加高频电场，对粒子进行加速。这种加速器被称

作同步加速器（synchrotron）。最早的同步加速器是加利福尼亚大学的麦克米伦建造的，于 1949 年开始运行，它可以在一个半径不到 1.3 米的圆形轨道上将电子加速到 320 兆电子伏的能量。

同步加速器存在的缺点是，一开始必须以某种方式向加速器中注入快速运动的粒子，被加速后的粒子必须被提取出来——这两个环节增加了仪器的复杂程度。从同步加速器得到的粒子还是成团的，要把粒子从入射速度加速到最高速度，就意味着加速频率和磁场强度都要从较小值慢慢变到较大值。

但它得到的能量却是巨大的！第一台质子同步加速器叫做宇宙线级加速器（Cosmotron）（图 4.5），于 1952 年在纽约的布鲁克海文国家实验室投入运行，它能够产生 3 吉电子伏的质子束。这至少 4 倍于大型同步回旋加速器所能达到的能量。重要的是，

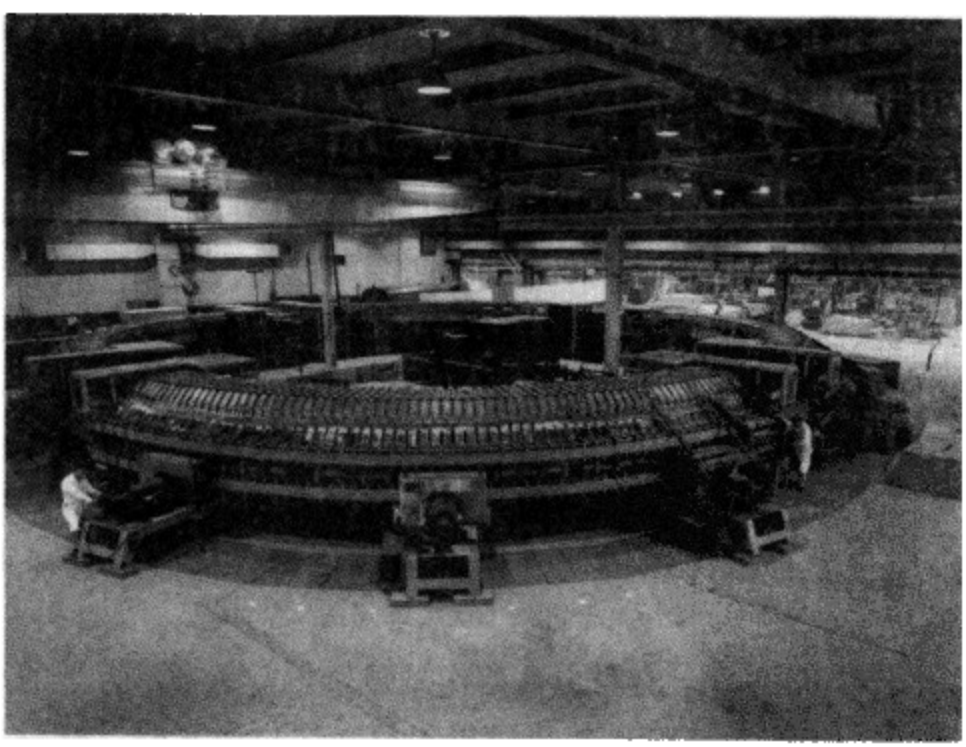


图 4.5　宇宙线级加速器 Cosmotron。布鲁克海文国家实验室的 Cosmotron 是世界上第一台质子同步加速器，于 1952 年投入使用。图片来源：布鲁克海文国家实验室

该能量足以产生像质子这样的重粒子，从此其他一些激动人心的新物质也开始陆续登场。

在 Cosmotron 开始运行的同一年，由欧内斯特·古朗特（Ernest Courant）、斯坦利·利文斯通（Stanley Livingston）和哈特兰德·斯奈德（Hartland Snyder）所组成的布鲁克海文三人小组，发表了一篇详细介绍控制粒子束方案的重要文献，称作交变梯度聚焦（alternating gradient focusing）方案，或者简称为 AG 聚焦。该方案用来解决原有加速器存在的重大难题，即粒子束会在中心路径的周围摇晃，粒子束的能量越高就越难控制这一问题。粒子束的晃动影响了微弱的聚焦效应——这种聚焦效应由将粒子限制在圆形轨道上的磁场产生，聚焦效应的破坏大大降低了粒子束的能量上限。

交变梯度聚焦是利用磁铁使粒子束稳定并产生汇聚的一种巧妙方法，其思想是利用粒子束所在的空间区域内不断变化的磁场——这就是磁场“梯度”。这种场可以用来使粒子束交替地汇聚或发散，从而形成一个整体上的总汇聚效应。实际上，四极磁铁就可以做到这一点。四极磁铁有两个南极和两个北极。给定一个四极磁铁，它的两对磁极能够在一个水平的平面内使粒子汇聚，而在竖直的平面内使粒子束发散。加速通道下游放置了另一个磁性与之相反的磁铁，从而会产生相反的效应。用这种方法来汇聚粒子束，看起来不太可能，但它确实有效。与之类似，我们可以用一个合适的透镜使光束汇聚或发散，从而出现净聚焦现象。

交变梯度聚焦，有时也被称作是强聚焦，康奈尔大学的小型电子同步加速器在 1954 年将该思想第一次付诸实验。第一台依赖于交变梯度聚焦的大型仪器是 CERN 的质子同步加速器，这台仪器在 1959 年启用。所有的现代环状加速器都用到简单的双极磁铁（称作偶极磁铁），通过均匀磁场约束粒子束，并且用四极

磁铁通过不均匀的磁场使粒子束汇聚。随着同步加速器原理提出以及聚焦问题的解决，又出现了其他一些阻碍大型加速器能量提高的问题，如经济上、政治上以及工程上的问题，这些都成为了限制加速器发展的因素。

现代粒子物理学离不开能将质子加速到几吉电子伏的加速器。只有在所有加速器运行的基本原理都搞清楚之后，这种加速器才能开始建造。如果没有考克饶夫和沃尔顿的同步加速器以及此后人们在这方面40多年的努力，这些原理、建造所需的技术以及政治、经济方面的投入就不可能存在。这些加速器出现之后，粒子物理才得以在地球上诞生。尽管现如今人们还在不断建造和使用小型回旋加速器，但在粒子物理学中，最引人注目的还是那些从小型加速器发展而来的造价昂贵的大型加速器。作为加速器的副产品，还诞生了许多相关的小设备——大约有15000台粒子加速器在全球各地的医院、电子工厂以及研究中心等努力地工作着。

随着粒子加速器的发展，对极小尺度的研究将需要投入极大的人力物力。要理解为什么需要建造如此巨大的加速器，并投入如此大量的资金是很容易的——大自然将四种基本相互作用力中的两种，都隐藏在了原子核当中，人们只能通过放射性衰变和一些人工嬗变来了解它们的性质。科学家们需要把一台足够强大的显微镜插入原子之中，用它来观察原子核的样子以及原子核的运行情况。这种显微镜就是粒子加速器，不断增加的能量会带来不断提高的放大倍数。有时仅仅靠观察是不够的，加速器还给人们带来了一种粉碎物质的方法，用来产生一些“拼图碎片”，并希望能利用它们拼出最终的答案。

它来自外太空

为什么我们一直说“地球上”的粒子物理学呢？实际上，大

自然早已非常体贴地为全宇宙的粒子物理学家们提供了一台天然粒子加速器。这是一台位于天空中的巨型粒子加速器。宇宙射线总是一直不断地轰击着地球，射线中大部分是在太空中被加速到很高能量的质子或其他原子核。尽管人们认为，被称作超新星（supernova）的恒星爆炸在宇宙射线的形成过程中起了重要作用，但宇宙射线的起源以及它们怎样被加速仍然是一个谜。人们只知道，当被称作初级射线的宇宙射线轰击地球大气层中的原子时，就会产生“次级射线”，也就是最终在海平面上探测到的那种粒子束。在到达海平面的宇宙射线中，大约有 3/4 是 μ 子（muon），其平均能量大约为 2 吉电子伏。在海平面上探测到的其他宇宙射线粒子还有中微子、电子和质子等。在地球表面上，每平方米每秒钟会有 180 个粒子通过，它们每时每刻都在轰击着我们。

某些宇宙射线具有高得不可思议的能量。位于犹他州沙漠中的高分辨率蝇眼探测器（HiRes，High Resolution Fly's Eye detector）和位于日本一座葡萄园中那台十分诱人的明野（Akeno）巨型天空阵列（AGASA），都探测到能量达到 10^8 太电子伏的奇怪射线。换句话说，一些轰击地球的射线，所具有的能量是现今世界上能量最高的加速器 Tevatron 所得到能量的 1 亿倍。

人们将宇宙射线看做是由天空中的巨型加速器产生的，这实际上是对这个科学谜团的一种揣测。宇宙射线从何而来？它们为什么会具有如此巨大的能量？是否它们诞生之时只是低能粒子，经过加速之后才具有如此大的能量？还是它们根本不需要加速，从某种过程中诞生时，本来就具有这么多能量？所有这些问题的答案没有人知道。

奥地利物理学家维克托·赫斯（Victor Hess）于 1912 年发现了宇宙射线，之后，实验物理学家们详细地研究了这些射线同物质的相互作用。他们经常在海拔比较高的地方进行实验，以便能获得能量更高的宇宙射线产物。他们所用的探测器就是简单的

照相底片，在底片上带电粒子的径迹表现为一条曝光的细线；他们有时也使用云室，通过水蒸气轨迹来追踪粒子。英国物理学家唐纳德·伯金斯（Donald Perkins）利用在 30 000 英尺（1 英尺约为 0.3 米）高空飞行的飞行器承载的照相感光乳剂，首次收集到了汤川秀树所预言的 π 子的迹象。他发现的证据就是一个小小的四角星形——一道只有百分之几厘米的微弱轨迹。同样是在 1947 年，由另一位英国物理学家塞西尔·鲍威尔（Cecil Powell）领导的小组证实了 π 子的存在，他们还了解到 π 子会衰变成 μ 子。鲍威尔的小组将照相感光乳剂分别放置在位于玻利维亚境内、安第斯山脉上海拔 5500 米的恰卡塔雅滑雪场（Chacaltaya），以及位于法国境内、比利牛斯山脉上海拔 2800 米的 Pic du Midi 天文台，让它们暴露在宇宙射线之下。

在宇宙射线实验中还发现了其他许多关键的粒子，包括正电子、μ 子和一种被称作 K 介子的粒子。20 世纪的 30 年代、40 年代和 50 年代初期是宇宙射线实验的辉煌时期。但宇宙射线是无法控制的，它们到达的时间和能量都无法控制。尽管最近对宇宙射线的研究开始慢慢复苏，但实验者们还是更愿意信赖建造加速器的工程师们。

传统上认为，1947 年 π 子的发现标志着粒子物理学的诞生。所有准备工作都就绪了。第二次世界大战之后，整个世界处于（相对）和平的时期，科学家们可以继续开展和平时期的研究工作。费曼、施温格、朝永振一郎以及其他一些物理学家们，都逼近了最终成为粒子物理理论模板的电动力学。有关 Cosmotron 和加州大学伯克利分校的十亿级加速器（Bevatron，用作转换粒子研究的质子加速器）的研究计划都在筹备当中。而且，那时的科学家们已经认识了 π 子，它不仅是原子核内部强相互作用的携带者，而且是除了光子之外，人们唯一知道的传递力的粒子。

像其他的新兴学科一样，粒子物理迅速发展了自己的术语

表，其中有一些词汇在 1947 年就已经存在，但大部分都是随后产生的。这些“新”词汇中的一些来源于希腊语，命名粒子的时候基本都同希腊字母保持一致，而且方程中的符号一般也以希腊字母表示。在有些情况下，希腊名称原来的意义被其他事件所取代，所以不需要追究希腊名称的原意，最简单的方法就是只记住它现在的定义。

现在我们来讨论一下介子，例如 π 子。所有介子的自旋值都是零或者某个正整数。这些介子不但参与强相互作用，并且还是强力的传递者。相互之间存在很强的相互作用，并且由介子传递相互作用的物质粒子被称作重子（baryon），例如质子和中子都是重子。所有的重子都受强力的作用，都很重，并且自旋都是像 1/2 或 5/2 这样的半整数。介子都是玻色子，重子都是费米子。

介子和重子统称为强子。反过来讲，可以简单地认为，一个强子就是任何一个参与强相互作用的粒子，它们也可能受到其他力的作用。例如带电的 π 子会受到电磁力的作用。从“强子”（hadron）一词可以衍生出形容词“强子的”（hadronic），例如强子碰撞（hadronic collision），它表示两个发生强相互作用的粒子之间的碰撞。

与之完全“相反”的是轻子（lepton），即那些不参与强相互作用的粒子。例如电子、μ 子和中微子都是轻子。所有已知的轻子自旋值都是 1/2，并且迄今为止它们都被认为是真正的基本粒子，也就是说它们不能被分割为更小的部分，也不存在内部结构。

到 20 世纪 50 年代初，刚刚起步的粒子物理学已经揭示，有三种发生强相互作用的粒子（即强子）存在，物理学家们希望通过它们来了解强相互作用。但他们还没有来得及对其中的规律进行了解，就马上又发现了许多其他粒子。

散射实验

物理学家们试图通过散射实验来了解强力的性质。最原始的强子-强子散射实验（即两个受强相互作用的粒子之间的散射）是质子-质子散射，实验的主要目的是测量散射截面，并将测量到的散射截面同理论计算值相比较，从而揭示原子核内部的规律。

涵盖内容最广的散射截面是总散射截面，许多核物理和粒子物理实验的目的都是测定总散射截面，它用 σ 表示，有时也记作 σ_T，其中 T 表示“总共”（total 的第一个字母）的意思。总散射截面标志着相互作用的总强度。粗略地讲，它代表碰撞粒子的几何尺寸。

例如在 1990 年，费米实验室的 E－710 合作小组在万亿电子伏加速器 Tevatron 上，测量了质子-反质子在 1.8 太电子伏的能量下发生碰撞的总散射截面，测量值为 72.1mb。mb 代表毫靶恩（millibarn），是一种粒子物理中表示散射截面值的标准单位：1 靶恩等于 10^{-28} 平方米，1 毫靶恩（即 1mb）等于 10^{-31} 平方米。一个面积为 72.1 毫靶恩的圆盘，其半径大约为 1.5×10^{-15} 米（即 1.5 飞米），所以它大约同质子的大小以及强力的范围相当。但总散射截面是随着能量的变化而变化的，因此用散射截面表示粒子的几何大小时应当特别小心。我们可以拿电子同质子的散射实验结果做一下比较，该实验测量到的质子半径大约为 0.8 飞米。

E－710 小组还测量了弹性散射截面。在测量弹性散射截面的实验中，两个碰撞粒子会互相弹开。他们测得的值为 16.6 毫靶恩，大约是总散射截面的 23%。总散射截面同弹性散射截面之差，就是非弹性散射截面，在上述例子中，它占总散射截面的 77%。非弹性散射截面表示了粒子的产生过程。

非弹性碰撞会产生非常壮观的奇异粒子喷射。在发生大量碰撞的情况下，每一次碰撞都会有粒子产生，所有产生的粒子又都各不相同，景象蔚为壮观。实验者们还会测量被他们称作“反应截面”（inclusive cross-section）的量。该过程可以用a+b→c+X来表示，其中 a 和 b 是发生碰撞的粒子，c 是实验者在寻找的特定粒子，X 是其他的所有东西。这样做的好处是，实验者可以设置他们的仪器，使其只对具有特定能量的 c 粒子实现“触发”。与反应截面的测量相反，测量“遍举截面”（exclusive cross-section）的实验可以表示为 a+b→c+d，在这种实验中可以测量和识别每个粒子。

实验物理学家发现强力

强子散射实验是强相互作用粒子之间的散射实验，最初的强子散射实验是卢瑟福、盖革（Geiger）、马斯顿（Marsden）和其他一些人在 20 世纪初所做的，即用 α 粒子轰击原子核的实验。但现代强子散射实验的前驱——类似于质子-质子散射之类的实验，直到 20 世纪 30 年代才出现。

但这明显是一个很原始的开端。从 1929 年到 1931 年，蒂宾根大学的克里斯蒂安·格思森（Christian Gerthsen）用最高达 100 千电子伏能量的质子，进行了质子-质子散射实验。在如此低的能量下，两个质子之间的最小距离大约为 10^{-13} 米；但这个距离太大了，根本无法观测到查德威克和比勒尔在 α 粒子-氢原子散射实验中所揭示的新作用力。实际上格思森只能验证，在这种能量下的质子-质子散射，只符合简单的静电力的预言。

而在华盛顿的卡内基研究院地磁部，布赖特和他的小组意识到，如果他们想要把质子加速到距“靶”质子足够近的距离，从而观测到除静电力之外的新物理现象，他们至少要达到 4 倍于格思森实验的能量。从 1926 年开始，布赖特就领导卡内基研究院

的一个小组，致力于用“高电压探索”来研究原子核的性质，由于那时还没有粒子加速器，这几乎是一个无法实现的要求。

作为该计划的一部分，第一次“高电压”质子-质子散射实验是由沃伦·韦尔斯（Warren Wells）于1935年发表的，其中“高电压”是一个十分主观的词。尽管粒子加速器已经出现了，但韦尔斯却没有使用它，他在实验中用到的高能质子是由钋源放出的α粒子轰击玻璃纸撞出来的。尽管质子的能量范围达到了1.5兆电子伏，但由于实验做得比较简陋，韦尔斯并没得到什么结果。韦尔斯查看了1.5万张立体相片，在20万条质子云室轨迹中，只发现了33次所谓“亲密接触的”质子-质子碰撞过程。

第二次尝试尽管也相当粗糙，但结果稍微好一些。那是在1934年，加利福尼亚大学的密尔顿·怀特（Milton White）用一台劳伦斯和列文斯通当初用过的小型加速器，将750千电子伏的质子射向氢原子靶。他们从拍摄下的大约25万条云室轨迹中，找到了约160次碰撞，1935年的《物理评论》（*Physical Review*）以简短的文字发表了这次实验结果。怀特写道：“尽管实验数据仍然不足，但足以看出它与质子是‘点电荷’的概念不符，这也足以写一篇让理论物理学家们很感兴趣的简单报告。”怀特的实验是世界上第一次用粒子加速器进行的质子-质子散射实验（图4.6和图4.7）。

卡内基研究院的默尔·塔弗（Merle Tuve）、诺曼·海登堡（Norman Heydenberg）和劳伦斯·哈夫斯塔德（Lawrence Hafstad）在1936年发表了一次更加精确的实验的结果。如果想要得到足够的数据，就需要在照片上找到数量更多的轨迹，实验的前景令塔弗和他的同事很失望；直到掌握了用电学方法为质子散射计数的技术之后，他们才能够继续进行他们的实验。与早期实验中产生的少量碰撞相比，他们用一台1.2兆电子伏的静电加速器，测量到成千上万拥有各种能量、不同角度的质子散射。他们

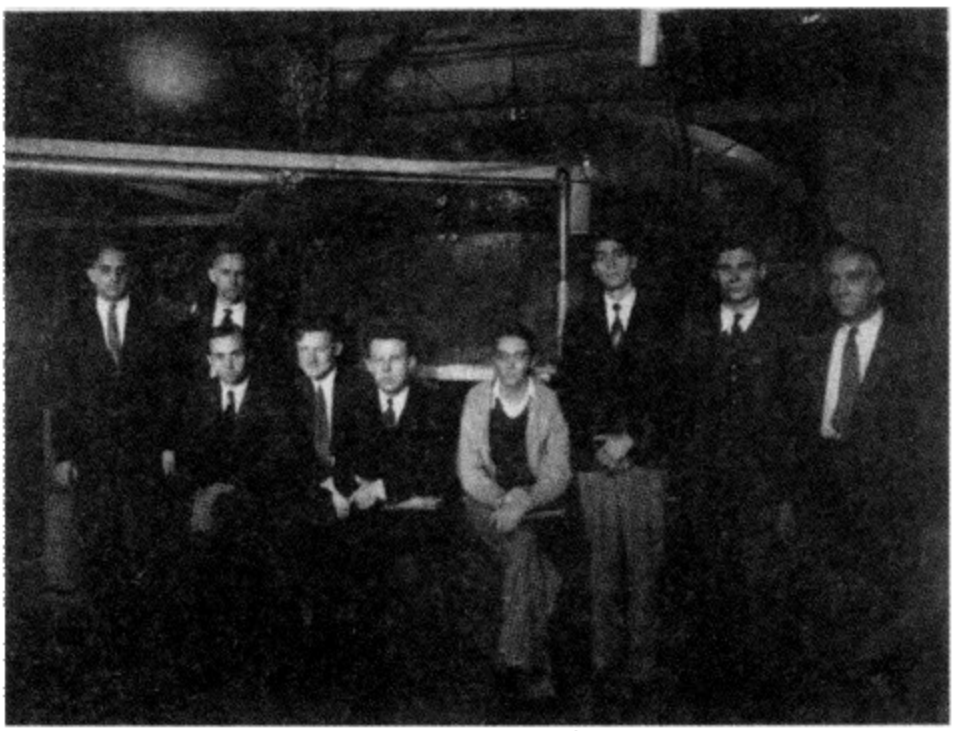


图 4.6　回旋加速器小组成员。欧内斯特·劳伦斯（正中）和他带领的回旋加速器小组，于 1933 年在 27 英寸的回旋加速器前所摄。斯坦利·利文斯通坐在最左边。密尔顿·怀特在劳伦斯的右侧，穿着一件浅色上衣。图片来源：劳伦斯-伯克利国家实验室

见证了简单静电散射的失效，在不到 5×10^{-15} 米的距离上，一种新的质子-质子相互作用出现了。

继在《物理评论》上发表那篇论文之后，他们又发表了由布赖特（Breit）、爱德华·康顿（Edward Condon）和理查德·布莱森特（Richard Present）合著的第二篇文章，这是一篇详细分析他们实验结果的文章。这篇文章将质子-质子散射实验的结果同中子-质子散射做了比较，尤其提到了位于罗马的费米小组的工作，最后得出了 p-p（质子-质子）相互作用与 n-p（中子-质子）相互作用完全一样的结论。该结论驳斥了原子核中的粒子是通过电磁力结合到一起的想法。并且，该结论也支持和鼓舞了电

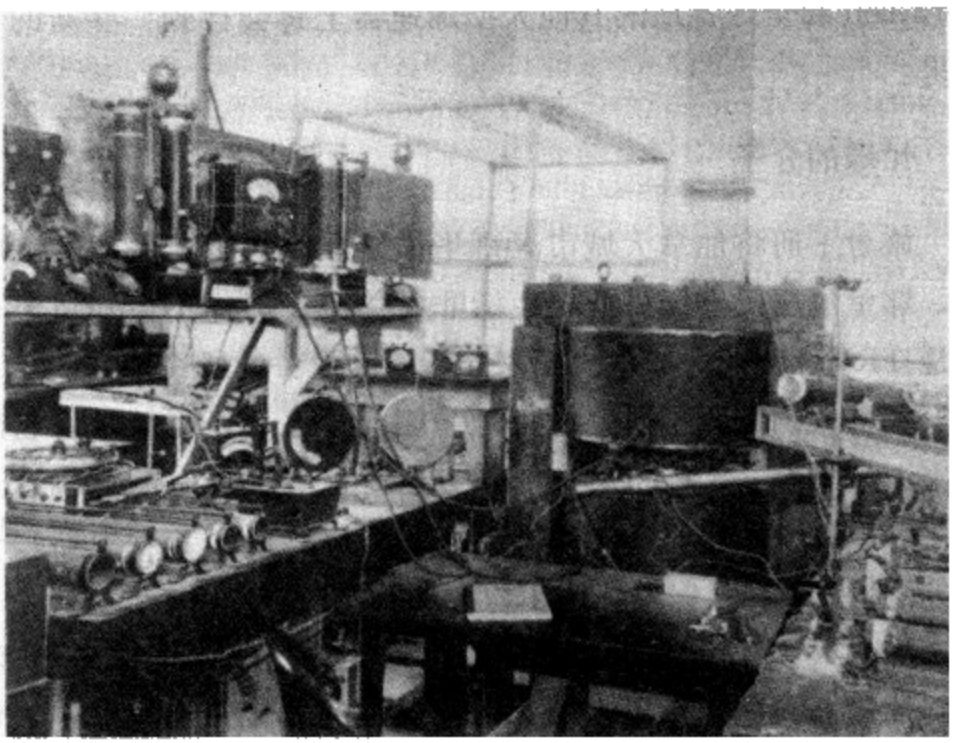


图4.7　11英寸回旋加速器，密尔顿通过它第一次完成了用质子轰击锂的实验，后来又利用它完成了高能加速器的质子-质子散射实验。这台加速器本身安放在一个大约28厘米见方，深约2.5厘米的黄铜“比萨盒”中。盒子被放置在一块大型磁铁的两极之间，照片中间的深色圆柱形物体就是大型磁铁。图片来源：劳伦斯-伯克利国家实验室

荷无关性想法，康顿和本尼迪克特·凯森（Benedict Cassen）运用海森伯在1932年提出的数学形式，给它穿上了现代数学的外衣。他们的论文在布赖特、康顿和布莱森特的论文发表之后也紧接着发表了。

因此，1936年对于强力来讲是至关重要的一年；这一年中，强力在实验上和理论上都获得了稳固的立足之处。从强相互作用角度看，质子和中子是完全一样的，强力要比静电力强得多，它将质子和中子紧紧结合在一起，这个过程与静电力无关。另外，一些早期的加速器实验提供了一些证据，这些证据预言在强大的

Tevatron 和全球各地的其他大型加速器上将会得到一些新的实验结果。

忙碌的介子

旅游手册将加拿大城市温哥华描绘成世界上最美丽的一座城市，那里的居民安居乐业，生活非常悠闲。但你不会知道温哥华还是 TRIUMF——加拿大国家介子研究设备的所在地。

加拿大国家介子研究设备是一座介子工厂，也就是说它能生产出大量具有一定能量的 π 子。它也是世界上仅有的三座介子工厂之一，另外两座是 LAMPF（洛斯阿拉莫斯介子发生器）和位于苏黎世的 PSI（保罗·谢尔研究所）。

加拿大国家介子研究设备的主要部分是一台直径为 18 米的回旋加速器，这是世界上最大的回旋加速器。在加速器中，带负电的氢离子（由氢原子结合一个额外的电子构成）被一块标志性磁铁引入机器内部的螺旋轨道，该磁铁由六部分组成，呈纸风车状。加速粒子每秒钟会受到 2300 万次电压“抽打”。在不到 1 秒钟的时间内，粒子速度就会增加到光速的 75%，并且达到最终所需的能量——520 兆电子伏。

粒子束会穿过三组金属箔片中的一组，金属箔片会将两个电子全部除去，只留下一个质子，质子就从加速器中飞出。因为条状的金属箔片可以在加速器内移动，因此粒子所得到的最终能量可以调节到 520 兆电子伏以下的任何值；三组独立的金属箔可以使一台机器产生三束能量不同的质子束。

2 号质子束被用于制造供医疗使用的放射性同位素；4 号质子束被引入进行质子碰撞实验的“质子厅”。而 1 号质子束也像前两束质子一样离开了加速器，被引入“介子厅”，在那里，质子束会打到碳靶或铍靶上产生出 π 介子。在十亿分之一秒的时间内，正负 π 子会发生湮灭，从而得到 μ 子和（反）中微子。而 μ

子本身在十亿分之二秒之后又会衰变成电子、一个中微子和一个反中微子。这束粒子的成分首先是质子，然后又变成 π 子，最后变成了 μ 子。各种粒子实验随着所用粒子的不同，在介子厅中也按照离加速器的距离从近到远排列。

在介子厅远端的尽头，科学家们可以利用已经被驯服的 π 子为癌症患者进行治疗。π 子放射线疗法是癌症治疗的一种形式，它可以攻击大脑或骨盆区域内的肿瘤。当一个 π 子通过一个活体组织时，几乎不会产生任何伤害，只有在马上要停下来时，它才会在那一点碎裂并放出射线，将周围的细胞杀死。通过这种方式，π 子像“深水炸弹”一样将放射线传递到癌细胞，而对其他组织几乎不会造成伤害。尽管作为人类与癌症抗争的一种武器，π 子在实践中并没有表现得像最初预料的那样出色，但它为放射线治疗提供了一种可选择的手段。

因为介子工厂主要是一个核物理研究中心，因此加拿大国家介子研究设备的大多数实验者还是将 π 子束指向一些没有生命的目标。许多实验致力于研究 π 子同质子或中子的散射过程。在加拿大国家介子研究设备产生的 π 子能量下，这些实验所测得的散射截面有一些与众不同的特征。特别是 π 子-质子散射截面，它一开始随着能量的增加迅速上升，在 π 子的动量为 195 兆电子伏时达到了最大值。随着能量继续增加，散射截面又开始下降，并且下降得同样很快，最终可以得到一个钟形的散射截面-能量图。

尽管按照现在的标准衡量，加拿大国家介子研究设备的 π 子束能量不是很大，但是在实验者们研究 π 子-质子散射的时候，他们所拥有的能量还是要比他们的前人大得多。在 1952 年，费米、赫伯特·安德森（Herbert Anderson）和其他一些人利用芝加哥大学的回旋加速器，首次观测到上文所提到的钟形峰的迹象，他们的 π 子最大可用能量只有 136 兆电子伏。大约在他们进行实验的 20 年之前，实验物理学家们用中子做最早的嬗变实验

时，就已经意识到了这种现象：这种峰是一种共振效应。

真正的瞬变现象

当某些系统的固有频率与外部对其施加的振动频率一致时，就会发生共振现象。当共振发生时，系统吸收能量，振动会变得特别强烈。比如，钢琴调琴师能通过确定音叉和琴弦之间的共振效应来调琴。在音叉“嗡嗡”震动的时候，把它放到琴弦旁边，然后调节琴弦的松紧，直到琴弦也一同振动起来。这时琴弦的松紧程度恰好可以发出正确的音调，它可以从那些由音叉传过来、经过它身边的声波中吸收能量。将一个扩音器放在一根线的上方，将线逐渐拉紧直到发生共振，然后继续拉紧，扩音器在这个过程中会出现短暂的音量下降现象。这种音量下降是由于在共振时，声音的能量被线所吸收，然后又朝四面八方再次发射出去；附近的人们听到的以及调琴师们利用的，正是这种再辐射的能量。

共振现象十分普遍，譬如广播的调台装置、管风琴的声音以及汽车所发出的令人讨厌的“嗡嗡”声都是由共振效应引起的。粒子加速器也利用了加速电压的频率和带电粒子绕行时间之间的共振。

太阳光中也存在共振现象。太阳光谱中各种颜色光的强度并不是平均分布的。可以让阳光通过棱镜将其展开，从而显示出它的色彩范围，这时你会看到阳光中有许多很细的暗线。这些暗线被称作夫琅和费线（Fraunhofer lines），是由于在太阳中心处所产生的某些特定频率的光，被太阳大气层吸收造成的。在太阳光谱的黄光部分中，被强烈吸收的钠 D 线就是一个经典案例，这说明太阳大气层中存在钠。

这些谱线对应着太阳大气层中原子能量的跃迁。当光的能量满足某个原子中一个电子从一个能级跃迁到另一个能级的要求

时，该原子就会将光吸收。由于原子再次将光放出的时候，光是朝着任意方向发射的，因此造成了从地球上看到的太阳光谱存在暗线。这就是在现实生活中，原子光谱出现的原因，它是原子量子本质的反映。

英国天文学家约瑟夫·洛克耶（Joseph Lockyer）通过研究太阳光中的暗线发现了氦。氦原子有两个电子，它的谱线比氢光谱复杂得多，但比其他更重的元素简单。但即使是很简单的氦光谱也存在许多奇怪的性质。

在人眼看不到的紫外波段仍存在氦谱线，但那里本来不应该有谱线。这些谱线的能量特别高，它所对应的光子能量，比将一个电子从氦原子中拉出来所需的能量还要大。出现这种明显的悖论要归咎于氦有两个电子。可见频段的光谱对应着其中一个电子的能级阶梯。而频率特别高的谱线与两个电子的共同作用相关。把两个电子一同提升到激发态的情况有可能发生，但要使这种现象产生，就要向氦原子中注入比能拉出一个电子的能量还要大的能量。氦原子会试图吸收更多的能量，但是这个过程很不稳定。最终不会形成一个处于激发态的氦原子，这种态迅速衰变为一个带正电的氦离子和一个电子。

这就是共振。射入氦原子的光符合产生激发态所需的能量，至少在理论上是这样。在实际中，光的能量太大了，这种激发态不稳定，最终发生衰变。

与之相反，在光谱低频区的谱线对应于氦原子在电子掉回基态前第二次吸收光子、电子跳到更高能级的行为，电子掉回基态时会发射光子。最重要的区别在于，这些态是束缚态（bound state），因为尽管电子处于激发态，但它们没有足够的能量逃离原子的边界。但在共振态中，能量足以让一个电子脱离原子，因此共振态不是一种束缚态。

典型的散射截面共振峰，是由于吸收光的频率处于或接近共

振频率造成的。这说明共振峰显示了频率上的一种散布，这种散布比“普通”光谱中所见到的谱线范围要大得多。这种散布反映了共振的不稳定性本质，是由某条能级受到它周围连续能级影响引起的。因此，处于一定频率范围内的能量都可以被吸收。根据不确定性原理，这种能量上的扩散（或叫做不确定性）可以理解为共振态的寿命。

当散射能量同某些不稳定的激发态相符，共振现象就会发生，结果会产生短暂的共振态。在原子物理、核物理以及粒子物理中都能够见到共振的身影。钢琴琴弦和管风琴管催生出了共振这个名词和共振理论，但这里所说的共振实际上是对能级而言的，是一种量子效应。

费米（图 4.8）是最先认识到 π 子-质子散射截面峰很可能是共振效应的几个人之一。在 20 世纪 30 年代中期，当时还在罗马大学的费米，成为研究中子散射尤其是低能中子散射方面的领路人。从中子诞生之日起，科学家们就认识到，可以把它作为探索原子核的理想工具。因为中子不带电荷，不受质子的静电排斥力影响，因此能轻易地到达原子核。费米与其他人一起，用许多元素研究了大量相互作用过程，为柏林的奥托·哈恩（Otto Hahn）和弗里兹·斯特拉斯曼（Fritz Strassman）开拓了道路。哈恩和斯特拉斯曼于 1938 年在被中子照射过的铀中发现了钡。他们的长期合作者利斯·迈特纳（Lise Meitner）那时已经流亡国外，她同她侄子奥托·弗里希（Otto Frisch）意识到，该发现标志着一个产生能量的新过程诞生了，他们将其命名为核裂变。哈恩获得了诺贝尔奖（化学奖），迈特纳却成为了政治和她人格的牺牲品，逐渐被人们所淡忘，但人类却因此得到了核能和原子弹。当然，这些成果都是之后发生的事情，当时费米和他的同事记录了许多中子被大量吸收的明显案例，他们推测这有可能是某种共振效应引起的。尼尔斯·玻尔（Neils Bohr）那时已经注意到，中

图4.8 意大利“航海家”费米，他的名字在“费米子”、“费米实验室”以及“镄”等名词中经常见到。费米利用中子引发核反应，并产生了新同位素。他也因此获得了1938年的诺贝尔物理学奖。费米在1942年12月2日首次实现了可控的自持核反应。那时和费米一同在芝加哥工作的亚瑟·康普顿，给曼哈顿计划管理委员会打去了一个电话，通报了这条新闻，他说：“意大利航海家已经登陆新大陆了。”图片来源：芝加哥大学

子通过原子核的时间大约是10^{-21}秒，而费米的实验却显示，中子在原子核周围耽搁的时间要比所需的时间长得多。玻尔受到这种现象的启发，在1936年提出复合原子核理论，该理论认为入射的中子进入原子核后，会将它的能量在所有组成原子核的核子之间进行分配。最终，处于激发态的不稳定原子核会发生衰变，可能会放出一个中子，或者是一个高能光子——即伽马射线（γ射线）。

玻尔的模型至今仍是描述原子核共振的一种实用图像。一次直接核反应所花费的时间，同一个粒子经过原子核的时间大致相同。而一个非束缚共振态会持续大约 10^{-18} 秒，大约是直接核反应时间的 1000 倍。与之相比，一个原子核束缚激发态会通过放出一个 γ 射线光子发生衰变，其寿命大约为 10^{-12} 秒，弱相互作用衰变过程会进行得更慢。

那时许多人都在考虑慢中子实验。在玻尔提出复合反应（等效于一种共振图像）的同一年，新泽西州普林斯顿高等研究院的布赖特和魏格纳，发表了一项在散射理论方面最为重要的结果。他们的论文题目十分简洁，叫做“慢中子俘获”。在这篇论文中，他们阐述了现在被称作布赖特-魏格纳公式的定理，该定理描述了散射截面共振曲线的特征形状。布赖特-魏格纳方法，将入射的物质波相位同散射后波的相位加以比较，波的相位受到散射中心的影响而改变。共振峰对应着两列波步调完全不一致的情况。布赖特-魏格纳的结论在实现中子散射十几年之前就已经提出来了，实际上它对于许多共振散射过程都适用。

粒子大爆发

在费米发表 π 子-质子散射结果 4 年之后，人们确认了这种峰是由共振现象引起的。实验者们发现，它们确实符合布赖特-魏格纳形式，它的寿命大约是 10^{-23} 秒——费米首次发现了粒子共振现象。

但是费米却不知道，他提出的 Δ 粒子共振是一场爆炸的第一声，这场爆炸就是现在说的粒子大爆发。尽管到 1950 年，一些“基本粒子”已经被发现，包括电子、正电子、质子、中子、中微子、μ 子、π 子以及 1944 年首次发现的 K 介子，但人们对接下来发生的事情还是没有任何准备。之后不久，人类在对宇宙射线的研究中发现了大量粒子，1951 年发现了质量为 1115.7 兆电子伏的中性

Λ 粒子，1953 年发现了质量为 1189.4 兆电子伏的 Σ^+ 粒子和质量为 1321.3 兆电子伏的 Ξ^- 粒子。

这一年也是加速器实验的势头超过宇宙射线实验的一年。世界上第一台质子同步加速器 Cosmotron 在一年前投入使用，并在 1953 年证实了 Σ^+ 粒子的存在，紧接着又发现了它的“姐妹”Ξ^- 粒子。在利用 Cosmotron 3 吉电子伏质子束所产生的高能 π 子的帮助下，人们证实了费米的 Δ（1232）共振态，并且发现它还伴随有包括 Δ（1920）和 N（1680）等的其他共振态——括号中的数字表示共振态的质量，也就是 π 子-质子系统在共振峰时的总结合能。

用同样的字母表示不同的共振态——比如 Δ（1232）和 Δ（1920）是一种近乎疯狂的想法。最新的数据表列出了包括这两种共振态在内的 22 种 Δ 共振态以及 21 种 N 共振态，其中 N（1680）按质量递增的顺序排在第 6 位。这确实是一次粒子大爆发。这种疯狂的想法认为，利用相同字母表示的共振态，尽管质量各不相同，但会通过某种方式相互联系。这些家族关系后来成为理解“基本粒子”动物园的核心内容。

大多数共振态都不像 Δ（1232）那样能够很容易地被认出来，Δ（1232）之所以很容易观测到，是因为它是整个大家庭中能量最低的共振态。能量较高的共振态会与同族成员以及其他家族成员发生重叠，从而在散射截面图上形成复杂的涨落，因此很难分析。

Δ（1232）共振态同它家族的其他成员一样，也是一种共振结构，也就是说是由相互碰撞的粒子结合形成共振态，之后它会发生衰变。另一方面，在共振态产生的过程中，会伴随产生一个或多个其他终态粒子。在这些情况下，会出现一个特殊的粒子集群，它们结合起来的总能量等于共振态的能量。实验者必须检测过程中产生的所有粒子，划分出所有可能的组合，以找到共振态

存在的证据，这使得这些实验很难进行。

但通过这种方法可以得到大量的结果，许多重要的共振态都是通过这种方式发现的。例如质量为 770 兆电子伏的 ρ 介子，于 1961 年被人们在 Cosmotron 上研究 π^- 粒子和质子碰撞时发现。碰撞会产生由 π 子与质子或中子组成的粒子对，这种粒子对在 π 子对的质量分布中会形成一个峰。同一年，搜寻共振态的能手之一——美国的路易斯·阿尔瓦雷茨（Luis Alvarez）和他的同事用质子加速器 Bevatron 找到了另一种重要的介子——质量为 781.9 兆电子伏的 ω 介子。他们研究了在质子-反质子碰撞过程中产生的三 π 子组团效应，该过程能产生五个 π 子。

大量新的共振态不断涌现：在 1960 年，实验者们发现了 Σ（1385）和 K*（892）。在 1962 年又发现了 Λ（1405）和质量为 547.3 兆电子伏的 η 介子。1962 年又发现了另一种重共振态 Σ（1530）和另外两种介子——质量为 1019.4 兆电子伏的 φ 介子和质量为 1275.0 兆电子伏的 f_2 介子。

共振态的典型寿命大约为 10^{-23} 秒，这使得它们成为人类研究过的最短命的现象。那么它们是真实的粒子吗？答案是肯定的，它们像质子和中子一样真实，但是有一点关键的不同。质子和中子都处在衰变链的尽头，它们不可能在满足各种各样守恒定律的情况下，再经由强相互作用衰变为其他粒子。但在本质上，共振态同质子和中子一样，都是“粒子”。

这也暗示了它们短暂的寿命。粒子物理中的共振态是由强相互作用创造并经强相互作用衰变的，这种衰变进行得特别快。如果这种由强力引起的衰变受到限制，比如受到某种守恒定律的限制，那么粒子可能就会变得稳定，或者以相对较慢的速度通过一些不违反守恒定律的其他相互作用发生衰变。例如中子和质子都是稳定粒子，都不发生由强力所引起的衰变。但是自由中子会衰变为质子（加一个电子和反中微子），其平均寿命将近 15 分钟，

这与共振态的寿命比起来，是一段相当长的时间。中子衰变之所以这样慢，是因为它的衰变通过弱相互作用发生，这说明由于不满足守恒定律，通过强力进行衰变的过程被禁止了。而最轻的重子——质子之所以稳定，是由于它不可能再衰变成其他更轻的重子。说质子稳定，实际上是指它们的平均寿命至少为 1.6×10^{25} 年；人们常常通过放置在矿井下的巨大探测器来检测质子的寿命，因为将强力、弱力、电磁力结合起来的“大统一理论”指出，质子实际上会发生衰变。但到目前为止，还没有人观测到质子发生衰变的现象。还有一些衰变是通过电磁力发生的，这些粒子的寿命大于共振态，但比通过弱相互作用衰变的粒子寿命要短，这反映了这三种力之间的相对强弱。例如中性 π 子经过电磁相互作用衰变为两个光子，其平均寿命约为 8×10^{-17} 秒；这比起发生弱力衰变的带电 π 子的寿命（2.6×10^{-8} 秒）要短得多。

如果没有粒子加速器，人们就不可能发现和研究共振态——它们的寿命太短暂了，人们不可能通过别的方式看到它们。对共振态的理解在 QCD 理论的发展过程中起到了重大作用。加拿大国家介子研究设备以及其他介子工厂，仍然在不懈地测量共振态的质量和衰变模式等性质，继续探索着共振态在低能过程中的各种特征。

这些新发现的共振态仅仅是“粒子大爆发”的一部分。狄拉克已经根据他的相对论性量子力学，预言了反粒子的存在——尤其是正电子的存在，正电子当时已经被发现了。但是，人们对同样的预言能否应用到像质子这样的重粒子上还存在怀疑，这在某种程度上是因为，如果他的预言成立，就说明会存在反物质乃至于反星系，而这些都是人们从未见过的。

“与质子相对应的反粒子是否存在”这个问题十分重要，因此人们运用质子加速器 Bevatron 巨大的能量来帮助我们搜寻这种尚存在于人们意识中的粒子。投入得到了回报。在 1955 年，由

两位美国物理学家欧文·张伯伦（Owen Chamberlain）和埃米利奥·西格里（Emilio Segrè）所带领的小组，在质子与铜靶撞击的碎片中发现了反质子。第二年，还是在 Bevatron 上，又紧接着发现了反中子。

随后，其他反重子也陆续地被发现，包括 1958 年发现的 $\overline{\Lambda}$、1960 年发现的 $\overline{\Sigma}$ 以及 1962 年发现的 $\overline{\Xi}$。到 1964 年，包括粒子和反粒子在内，所有粒子的总数超过了 60 多种。单从数字上也可以毫无疑问地看出来，其中并没有几种（如果有的话）粒子是真正的基本粒子，这种想法也得到粒子所表现出的规律性的支持。物理学家们知道，一定存在一种方法来理解和管理这么多纷繁复杂的粒子。最终，他们的确找到了一套解决方案，将发生强相互作用粒子的家谱梳理得井井有条，这套方案的基础是一些构成物质的新基本“砖块”。这些砖块有一个名字，叫做夸克（quark）。

八正法

物理课本和科普读物本身都存在一个内在的问题，那就是认为物理学上的“伟大思想”不容置疑。问题在于，在这些书中有许多结论看起来是经过精妙推理得出的，但却不说明从何而来，这些结论立即就成为了“伟大思想”，被所有人通通接受。这就会给人们以这样的印象：当时所遇到的所有紧迫问题都已经得到解决，研究者们只需要准备好，等着从逻辑上对其进行下一步的梳理就行了。

QCD 理论的故事会反复告诫我们，通常的情况都不会是这样。但无论专业文章，还是通俗的物理学报道，都会砍掉不确定的内容、误解、错误的假设、笨拙的标记以及大量客观资料，以避免学生和不太了解的读者陷入困惑和无所适从当中，但这些都是“伟大思想”的基础。这是一个可悲的现实，因为它展现给人

们的是一个“清扫过”的整洁的物理，与混乱的现实相去甚远。

夸克就是一个伟大的思想，是最伟大的思想之一。夸克还为我们提供了一个伟大的思想是怎样发展起来的范例，并且说明这些伟大思想在面对整个问题时通常会怎样的无能为力（但至少在一开始，人们提出这些想法的目的就是要解决这些问题）以及其他研究人员怎样将它们看作是疯狂或错误的想法。这就是从不同角度、从QCD理论的众多发展阶段以及整个物理学上对它的评价。夸克的故事是一本展现真实物理规律的生动教材。

宇宙之谜

夸克的故事可能真正开始于由英国曼彻斯特大学的乔治·罗彻斯特（George Rochester）和克利福德·巴特勒（Clifford Butler）制出的一张模糊的灰度图。他们在一张布满5000条宇宙射线轨迹的云室照片中，找到了2个“具有惊人特征的分岔轨迹”，引号中的句子引自他们1947年在《自然》（*Nature*）杂志上发表的一篇报告。

罗彻斯特和巴特勒相信，他们已经找到了发现新基本粒子的证据。根据他们的分析，有一张照片显示的是中性粒子衰变为两个带电粒子的过程，该中性粒子的质量为340～800兆电子伏。另一张照片则显示了一种带电粒子的衰变过程，它的质量为490～940兆电子伏，但它不是质子，这种带电粒子能够衰变为另一个带电粒子和一个中性粒子。在解释这两个过程时，供他们选择的粒子只有核子、质量为105兆电子伏的μ子以及质量大约是140兆电子伏的π子，因此很容易看出，他们为什么得出了发现新粒子的结论。

这些粒子轨迹十分令人吃惊，由此得出的结论也令人激动，但这在当时却没有引起人们足够的重视。两年后，加州理工学院（Caltech）的卡尔·安德森（Carl Anderson）和他的小组也在他

们的云室照片中找到了类似的粒子轨迹，在曼彻斯特大学得出的结果被证实了。曼彻斯特小组和世界上其他一些宇宙射线小组都做了进一步的实验，最终发现了两种新的中性粒子，即如今所知的质量为 1115.7 兆电子伏的 Λ 粒子和质量为 597.7 兆电子伏的 K^0 粒子。早期的实验还显示，Λ 粒子能衰变为质子和某种未知的粒子，现在人们已经知道，这种未知粒子就是带负电的 π 子。K^0 粒子也有带电的“同伴”，现在被称作 K^+ 和 K^-。实际上，人们在 1944 年的某一天第一次观测到 K^+ 粒子，在法国阿尔卑斯山地区工作的路易斯（Louis Leprince-Ringuet）和迈克尔（Michel L'héritier）发现了一条单独的轨迹，但现在人们对他们当时发表的结果尚持怀疑态度。人们观测发现，K^+ 粒子最终会衰变为 μ^+ 粒子和其他的未知粒子——现在知道这种未知粒子就是中微子，这很可能就是罗彻斯特和巴特勒在其中一张照片中所观测到的那种带电粒子衰变。人们认为，他们观测到的另一种衰变实际上是 $K^0 \rightarrow \pi^+ + \pi^-$ 过程。

表 4.1 列出了那个时代人们所知的粒子家族中寥寥无几的几个成员，随着新粒子的发现，这张表不断被扩充。人们对新粒子了解很少——数据支离破碎，进行统计的工作十分艰难，在测量

表 4.1　　在 1952 年左右人们已知的粒子

粒子	质量（兆电子伏）		粒子	质量（兆电子伏）	
γ	0	光子	π^0	135.0	π 子
e^-	0.51	电子	π^+	139.6	
e^+	0.51	正电子	π^-	139.6	
μ^-	105.7	μ 子	p	938.3	质子
μ^+	105.7	反 μ 子	n	939.6	中子
K^0	497.7	奇异介子	Λ	1115.7	奇异重子
K^+	493.7		Δ	1231.0	共振态
K^-	493.7				

质量和估计散射截面的时候，大量云室轨迹都靠不住。得到的结果也令人伤透了脑筋，因为很难辨别两次不同的衰变是源自于两种不同的粒子，还是来自于同一种粒子的两种不同的衰变模式。没有一种衰变事件会产生两种完全相同的衰变产物。

但是有一件事情很清楚：新粒子给人们带来了一些难题。新粒子在变为其他粒子的过程中留下了一条很长的轨迹，由此可以看出衰变过程大约经过了 10^{-10} 秒。这应该是弱相互作用衰变所需要的时间，比强相互作用衰变的时间要长得多，因此说明衰变通过弱力进行。并且存在足够的数据说明，这些粒子产生的频率与 π 子差不多，这又说明它们由强相互作用产生。但如果是这样的话，它们为什么不通过强力进行衰变呢？像费米新发现的 π 子-质子共振态那样由强相互作用产生的粒子，一定要以同样的方式进行衰变吗？其中间态的寿命一定大约是 10^{-23} 秒而不是 10^{-10} 秒吗？基本问题在于如果假设新粒子是通过强力产生的，它们的寿命为什么会这么长。正因为具有这种奇怪的行为，这些粒子被统称为“奇异粒子”（strange）。

在不存在任何描述核子相互作用的可靠理论的指导下，普林斯顿高等研究院的亚伯拉罕·派斯（Abraham Pais）还是利用有限的数据，于 1952 年在解决这些混乱的问题上取得了进展。派斯认真思考了通过强力诞生又通过弱力衰变的过程，提出一套简单的数字规则，其中强相互作用（以及电磁作用）满足一套规则，而弱相互作用满足另一套不同的规则。按照要求，他的规定禁止了奇异粒子通过强相互作用发生衰变，并且能够预言奇异粒子会成对地产生。

这看起来也许不怎么样。实际上，派斯仅仅把他的想法看做是“努力将现有的信息归类整理的方式”，即他所研究的是一种分类定理。他的规则是试图为神秘而寥寥无几的实验结果赋予意义的一种工具，并且它也做到了这一点。它甚至还提出了一个预

言，事实证明，这个预言至少部分正确。但其欠缺之处在于，它还不是一套无懈可击的完整理论，并且它也不是独一无二的，因为南部阳一郎（Yoichiro Nambu）、西岛和彦（Kazuhiko Nishijima）、大根田治男（Sadao Oneda）和其他一些人都进行了相似的推理过程。但是他的想法指明了前进的道路。简言之，派斯的论文是真正的科学。

那么，派斯的那个预言到底怎么样了呢？就像派斯预言的一样，实验者们在质子同步加速器 Cosmotron 上见证了奇异粒子成对产生的过程。为了完成这项实验，他们将 1.5 吉电子伏的 π 子束射向充满氢的云室，得到了反应 $\pi^+ + p \rightarrow \Lambda + K^0$：对于一台加速器来讲，产生奇异粒子是十分简单的事情。实际上 Cosmotron 和 Bevatron 都能够大量产生奇异粒子，在人们的控制下，产量能达到每天上千，甚至上百万。研究宇宙射线的科学家们可能要花几周甚至几个月的时间，从成千上万的照片中间找出偶尔存在的奇异粒子，这是绝不可能完成的事情。加速器宣告了研究宇宙射线的经典时代已经终结，同时也为人们提供了探索新奇异粒子的方法。它们马上就扩展了粒子大家庭成员的名单：Σ^+ 和 Σ^- 都是很重的奇异粒子。随着加速器的发展，事情的进展速度非常快。

在另一种重奇异粒子——Ξ^- 被发现之前，《物理评论》在 1953 年发表了一篇只有半页的文章，这篇文章为提出奇异粒子所遵循的规则体系创造了条件。这种新的方案十分重要。论文的作者是芝加哥大学的一名讲师，他的名字叫做默里·盖尔曼（Murray Gell-Mann，图 4.9）。

奇异对称性

如果想找出一个既与夸克有关，又与 QCD 理论相关联的特殊人物，那就非默里·盖尔曼莫属。但是当盖尔曼挑战强产生、弱衰变的奇异粒子问题时，夸克和 QCD 理论恰好都步入了正轨。

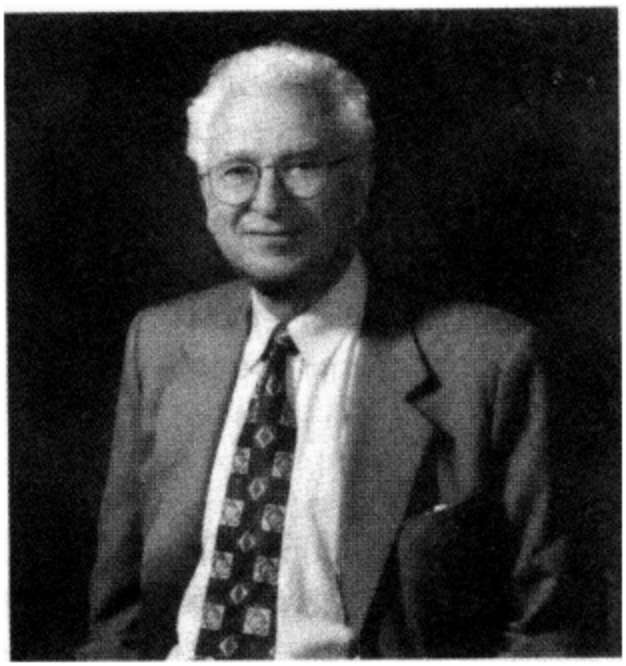


图 4.9 盖尔曼。在 QCD 理论的故事中，再没有别的人名比他的名字出现的次数更多了。图片来源：Louis Fabian Bachrach

盖尔曼受到了派斯和哥伦比亚大学的物理学家大卫·比斯里（David Peaslee）的启发。后者曾提出新粒子也应该遵从电荷无关性原理，并且开创性地使用了同位旋理论——盖尔曼非常支持该想法。他假设 Σ^+、Σ^0 和 Σ^- 三种粒子共同组成了一个同位旋家族，它们作为一个家庭中的成员都具有相同的同位旋值，Σ 粒子家族成员的同位旋值都是 1。另外，他还指出 K^+ 和 K^0 粒子会形成另一个同位旋为 1/2 的家族，而 $\overline{K}^0$ 和 K^- 粒子也组成了同位旋为 1/2 的家族。

这下可完了，盖尔曼的短文包含了一些相当疯狂的想法。他的粒子族群中包括 Σ^0 粒子，但它那时还没有被发现——直到 1955 年，人们才发现了 Σ^0 粒子。此外，盖尔曼的家谱中并不包含当时已经存在的 Λ 粒子。还有，他的理论在 K 介子家族的分配上也不再那么简单了，它要求与 K^0 粒子相对应的反粒子不再是

K^0 粒子，而是 K^- 粒子；并且与 K^+ 相对应的反粒子是 $\overline{K}^0$。这使一些人感到吃惊和不舒服，因为他们觉得光子的反粒子是它自己，即 $\overline{\gamma}$ 等于 γ；而 π^0 介子也是它自己的反粒子，即 $\overline{\pi}^0$ 等于 π^0，那么 K^0 也应当遵从这套规定。但是没有任何证据证明它一定会怎样，盖尔曼突破了这种常规的看法。他的想法遇到了一些反对，但是在同一年，大阪城市大学的西岛和彦、中野董夫也独立地得到了同样的结论，盖尔曼的理论获得了应得的支持。

在随后的两年中，盖尔曼与西岛和彦各自修正了自己的粒子分类理论，推演出基于物质的一种全新性质——“奇异性”(strangeness) 的两套完全相同的理论。奇异性是一种简单可加的守恒量子数，某个特定同位旋家族中的粒子，都具有同样的奇异量子数。根据这套规则，三种 Σ 粒子的奇异量子数值都表示为 $S=-1$。而 K^+ 粒子和 K^0 粒子的奇异量子数 $S=1$，$\overline{K}^0$ 粒子和 K^- 粒子的奇异量子数 $S=-1$。质子和中子不具有奇异性，因此它们的奇异量子数 $S=0$。这种理论认为，在强相互作用过程中，作用前后奇异量子数的总和必须相同，但是弱相互作用过程就不必满足这个条件。例如 Λ 粒子的奇异量子数 $S=-1$，奇异数理论禁止中子-中子相互作用过程 $n+n\rightarrow\Lambda+\Lambda$，因为左边的奇异数总和为零，而右边的奇异数总和为 -2，但派斯的理论却允许这个过程发生。而反应 $\pi^++p\rightarrow{}'\Sigma^0+K^+$ 却是一种可行的强相互作用过程，因为奇异数的分配情况为 $0+0=-1+1$。但是衰变过程 $\Lambda\rightarrow p+\pi^-$ 不可能是一种强力衰变过程，因为奇异数分配为 $-1=0+0$，这个式子是错的，因此衰变是由弱相互作用控制的。

引入一种全新的物质性质以及一种新的守恒量子数，并不是物理学家们每天都会做的事情。但是电荷守恒定律以及同位旋守恒定律已经开了先河。此外物质还存在另一种简单可加的守恒性质，这种性质与物质宇宙的实际稳定性有关。

如果不考虑变老的情况，每个人都知道我们不会慢慢地发生

分解。我们赖以生存的物质世界也不会发生分解。这是质子不发生衰变的标志。与之相反，中子很乐意在几分钟之内衰变为质子或其他某种粒子。但即使是中子，也会区别对待它们可能进行的衰变，例如它会拒绝只衰变为 π 子。这一切都提醒我们，应该存在某种有关核子的守恒量子数。该量子数被称作重子数，记作 B。

质子和中子的重子数都是 1，反质子和反中子的重子数为 -1。迄今为止，在人们观测到的所有反应中，重子数都守恒。例如中子的弱衰变过程 $n \to p + e^- + \bar{\nu}$ 在该理论下是允许发生的，因为重子数分配情况是 $+1 = +1 + 0 + 0$。质子不存在相应的衰变，是因为质子比中子稍微轻一些，因此无法通过弱相互作用衰变为其他重子，它已经处于这条衰变链的最末端。像 $n \to \pi^+ + e^-$ 和 $p \to \pi^0 + e^+$ 这样的衰变过程是不允许发生的，因为这两个过程的重子数分布都是 $+1 = 0 + 0$，是错误的。重子是发生强相互作用的物质粒子，它们至少和质子一样重，它们的重子数 $B = 1$。反重子的重子数都为 $B = -1$。

重子数守恒是一种经验法则，但也得到量子场论的支持，核子在场论中可以被看做是像电子一样的基本费米子。那么既然电子有与其对应的反粒子，核子也一样——因此存在反质子和反中子。但量子场论要求粒子数——正粒子的数量减去反粒子的数量也是守恒的。在 QED 理论中，粒子数就是电子数减去反电子数，这种“守恒荷”可以用电荷来标识。对于核子来讲，核子数减去反核子数也要守恒，相应的守恒核过去被称作“重荷”，实际上就是重子数 B。重子数 B 守恒等价于核子主方程满足全局 U（1）对称性；这就像对电子和正电子来讲，全局 U（1）对称等价于全局电荷守恒一样。但在电荷的情况中，U（1）对称和电荷守恒都要变成局域的，因此会出现相应的传递力的粒子——光子。对于重子数 U（1）群却没有对应的步骤，这表示它不会导出一套可用的粒子力理论。

实际上还存在轻子数，轻子数满足的规律与重子数的相似。产生电子的时候也会伴随产生一个相应的正电子，但是如果只需要考虑电荷守恒，那么同时产生一个 μ^+ 粒子来代替正电子也是允许的。我们先看看其他例子，像 $e^+ + e^- \rightarrow \mu^+ + e^-$ 这样的过程，尽管也满足电荷守恒，但在自然界中不允许出现，而过程 $e^+ + e^- \rightarrow \mu^+ + \mu^-$ 则不存在任何问题。根据经验规律，规定电子和与电子相关的中微子所具有的轻子数为 $L_e = +1$，规定它们反粒子的轻子数为 $L_e = -1$。μ 子以及 μ 子中微子（与电子中微子不同）所具有的轻子数为 $L_\mu = +1$。在这里，下标中的 μ 和 e 强调了 μ 子中微子和电子中微子的区别。μ 子以及与 μ 子中微子的反粒子所具有的轻子数为 $L_\mu = -1$。根据这套理论，早期宇宙射线中的 K^+ 衰变过程实际上是 $K^+ \rightarrow \mu^+ + \nu_\mu$，$\mu$ 轻子数分配为 $0 = -1 + 1$，其中 ν_μ 表示 μ 子中微子，其轻子数是 $+1$。第三个轻子数是 L_τ，标记另一种轻子 τ 子（tau，音“套”）。马丁·佩尔（Martin Perl）和他的小组于 1975 年在 SLAC 发现了很重的 τ 轻子。佩尔的发现使他获得了 1995 年的诺贝尔物理学奖，莱因斯（Reines）由于发现电子中微子而与他分享了这一奖项。τ 子也有与之相对应的中微子，直到 1998 年，费米实验室才首次直接观测到了 τ 子中微子，它们的轻子数 L_τ 的值都是 $+1$。像重子数一样，在所有反应中各种轻子数总是满足各自的守恒定律；并且它们也同重子数一样，与相应场的全局 U（1）对称性相关。

目前仍存在一些难以理解之处。人们认为纯粹的全局对称性是偶然出现的，在基于局域对称性的规范理论中并不重要。是不是它们根本就不是纯粹的对称性？也就是说，相应的“守恒量”可能并不是真的守恒？1974 年，同在哈佛大学工作的霍华德·乔吉（Howard Georgi）和谢尔登·格拉肖（Sheldon Glashow）引入了将强力、弱力和电磁力都统一到一个框架中的理论，至今这仍是最简单的理论。他们的方法是格拉肖、温伯格和萨拉姆的电

弱相互作用理论在逻辑上的扩展，是基于一个更大的对称群——SU（5）群的。他们的大统一理论最惊人的预言，是 B 和 L 各自不独立守恒，而 B、L 加起来则守恒。如果他们是正确的，那么物质实际上是不稳定的，我们以及我们周围的一切都在慢慢地衰变。人们做过一些观察质子衰变的实验，最典型的实验是在矿井坑道中放置大量物质，从而观测质子偶然发生的衰变。在矿井里，宇宙射线所造成的破坏会降到最低。但至今尚未观测到任何质子衰变的证据。因此即使宇宙在发生衰变，也进行得特别缓慢。根据乔吉和格拉肖理论的一个简单推论，质子衰变最可能是通过 $p \rightarrow \pi^0 + e^+$ 过程实现的——这明显违反简单的重子数和轻子数守恒，质子的平均寿命应该低于 2×10^{31} 年。由于最近的实验证明质子寿命不低于 9×10^{32} 年，如果按照这一衰变过程，质子要比“简单”理论预言的稳定得多。因此我们不得不放弃乔吉和格拉肖的基本理论。

但这些问题的出现，已经是 20 年之后的事情了，当时盖尔曼和其他人还在试图理解强相互作用的基本规则。重子数和奇异量子数最终成为了非常关键的量子数，同位旋也是如此。

有什么样的自旋，就有什么样的同位旋

同位旋是模仿自旋概念提出来的。一种将自旋与“常见物体”类比的方法是，考虑一个绕固定点转动的刚体杆，杆的一端连接在固定点上。这种转动形成一个 SO（3）群。将自旋概念应用于电子这样的粒子上，恰当的对称性描述是 SU（2）群，它与 SO（3）群只有细微的不同。与一个旋转的陀螺要么顺时针转动，要么逆时针转动类似，电子的自旋值也有限制，这可以从某些原子光谱的分裂中看出来。

同位旋同自旋很相似，它有两个核子态（一个是中子，另一个是质子），对应着两个自旋态（向上和向下）。同位旋“转动”

对应于某种抽象的“同位旋空间”，而不是某种真实空间——同位旋是一种内在性质，而不是时空变换。但用于自旋的所有数学都可以直接应用于同位旋，见图 4.10 的图示说明。

质子和中子都属于同一个“同位旋双重态”，双重态表示它们是两个具有某种共同性质的东西，在该例子中这两种东西都是核子。关键之处在于，强相互作用在两个可能的同位旋态之间没有任何不同——强力只知道它们都是核子。这两个核子态分别表示为：$I_3=+1/2$ 的核子是质子，$I_3=-1/2$ 的是中子。注意下标中的数字 3。下标 1、2、3 表示同位旋 I 的 3 个分量，就像在讨论自旋或转动时人们习惯用到的 x、y、z 一样。之所以用数字，是为了强调同位旋空间是一种内禀空间，而不是粒子所在的物理空间。

除质子和中子外，其他粒子也可以组成同位旋家族。3 种 π 子——π^+（139.6 兆电子伏）、π^0（135.0 兆电子伏）和 π^-（139.6 兆电子伏）——尽管它们所具有的电磁性质将它们划分为 3 种不同的粒子，但只要涉及与强力有关的性质，它们三个是完全相同的。它们组成了同位旋三重态，或者说是一个三口之家——一个具有三种可能同位旋态的 π 子家族。如果“派”（π）是该家族的姓，其同位旋 $I=1$，那么三个家庭成员的名字分别是“派加”（π^+，$I_3=+1$）、“派零”（π^0，$I_3=0$）和“派减”（π^-，$I_3=-1$）。还存在四口之家，也就是同位旋四重态的情况，如 Δ（1232）共振态家族，同位旋为 $I=3/2$，其 I_3 值分别为 $+3/2$（Δ^{++}）、$+1/2$（Δ^+）、$-1/2$（Δ^0）和 $-3/2$（Δ^-）。

物理学家们能够用同位旋预言与强相互作用相关的散射截面。这样做就意味着要同时考虑粒子和它们的同位旋。其原型仍然是自旋。

自旋是一种向量，其性质不仅与它的大小（“自旋箭头”的长度）有关，还与它所指的方向有关。向量叠加的规则必须既考虑

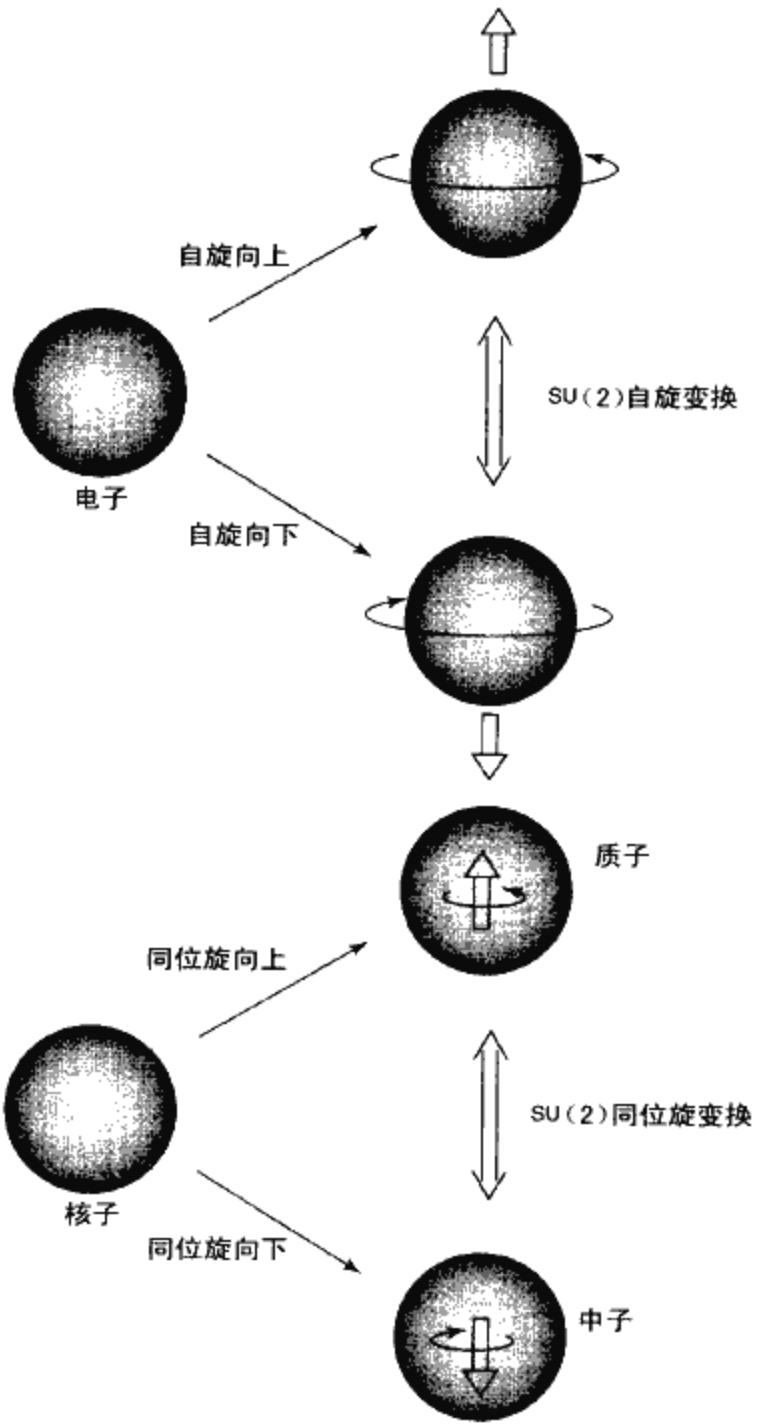


图4.10 将自旋的数学描述用于同位旋，图中的“转动”都位于内禀空间内

向量大小，又要考虑向量的方向信息；像电子自旋这样的“量子向量”，还必须满足向量可能取值的限制。对自旋为 1/2 的电子来讲，自旋箭头的长度为 1/2，其方向要么指向上方，要么指向下方。两个这样的粒子能够组成一个自旋值为 1 或 0 的结合体。但是由于每个粒子都有两个可能的态——向上或向下，就存在 4 种可能的组合。如果两个粒子的自旋都向上，那么总自旋的 z 方向分量值将会是 $+1$。如果两个都朝下，那么总自旋的 z 方向分量值为 -1，见图 4.11。在这两种情况下，其结合体总自旋的 z 方向分量很明显是 1。两个粒子的自旋也可能相反，这样总自旋的 z 方向分量就是 0。

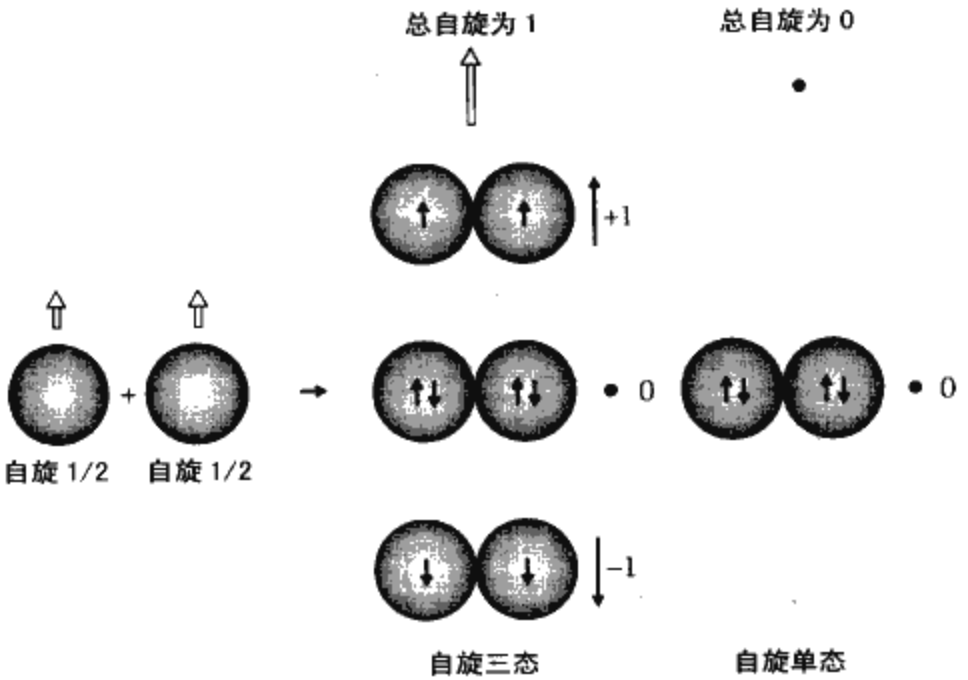


图 4.11 自旋（或同位旋）合成。两个自旋为 1/2 的态（自旋向上或向下的态）合成，得到 3 种总自旋为 1 的态。这 3 种态自旋的第三方向分量分别为 +1、0 或 −1。此外，还存在一种总自旋为 0 的态

还有一些东西没有说清楚。这里只有三种情况，第四种情况在哪儿呢？从 z 方向来看，自旋方向相反的情形包括两种情况：

第一个粒子向上、第二个向下，或者是第一个向下、第二个向上。一个总自旋为1的结合体从自旋可能性的角度来看，就像一个“基本的”自旋为1的物体一样。按照法则，这种自旋为1的物体有3种可能的自旋z方向分量：+1、-1和0。

最后一种可能性看起来像失踪了——它为我们提供了一种处理自旋叠加问题的法则。将两个粒子A和B的自旋以{A（上）B（下）+B（上）A（下）}的方式叠加，会得到自旋z方向分量为0的结合体。以{A（上）B（下）-B（上）A（下）}的方式叠加，也会得到同样的结果。这两种组合方式看起来很相似，但实际上有很大差别。想要看出这种差别，一种方法就是将粒子A和粒子B对调。换句话说，就是交换粒子的标签而保留自旋分配不变，也就是将A（上）变为B（上），将B（下）变为A（下），等等。因此，第一种组合方式就变成了{B（上）A（下）+A（上）B（下）}，同变化前的样子完全一样。但是第二种组合方式变成了{B（上）A（下）-A（上）B（下）}，其实也就是负的{A（上）B（下）-B（上）A（下）}，与原来的状态完全相反。

看到两个粒子组成的结合体在两个粒子互相交换时的情形，我们会有一种似曾相识的感觉——我们可以回想一下两个电子组合所具有的交换反对称性。对于自旋来讲，{A（上）B（下）+B（上）A（下）}的结合方式满足交换粒子标签A和B操作下的对称性，而{A（上）B（下）-B（上）A（下）}由于在交换标签A、B后，整体上会出现一个负号，因此满足反对称性。这种组合方式相对于其他三种方式显得有些与众不同，因为在A（上）B（上）和A（下）B（下）这两种自旋z方向分量分别为+1、-1的情况下，交换A、B粒子的操作也满足对称性。因此，它成为了唯一满足反对称性并且自旋值为零的“孤独的”组合方式，因为它的总自旋值为0，自旋在z方向的分量值同样也是0。

而混合组合方式｛A（上）B（下）＋B（上）A（下）｝的自旋值也是1，只是自旋的 z 方向分量为0。通过这种方法，3种自旋为1的态具有相同的对称性。这种对称性同自旋为0的混合方式所具有的对称性完全相反。

由于同位旋与作为原型的自旋极其相似，两个同位旋为1/2的粒子相叠加，会得到总同位旋为1，同位旋第三分量为＋1、0、－1的三个态，以及一个总同位旋为0的态，其第三分量也为0。

同位旋为1/2的粒子组合方式，可以用来理解质子和中子的散射。但从历史上看，使同位旋声名鹊起的还是 π 子和核子之间的散射过程。像核子与 π 子这样同位旋为1/2的粒子与同位旋为1的粒子组合，得到的总同位旋可能为3/2和1/2。例如负 π 子同质子发生碰撞可以产生叫做N（1520）的共振态，它的同位旋值为1/2（图4.12）。一般的过程在反应前后只满足 I_3 值的总和相等——I 值要满足更复杂的向量叠加规律。另一个例子是正 π 子同质子发生碰撞生成 Δ^{++} 共振态的过程，其同位旋值 $I=3/2$，并且 $I_3=3/2$。强相互作用不依赖于 I_3 的值，因此 $\pi^+ + p \rightarrow \pi^+ + p$ 过程的散射截面同 $\pi^- + n \rightarrow \pi^+ + n$ 过程的散射截面是一样的，在过程中只有 $I=3/2$ 起作用。但强相互作用确实依赖于总同位旋 I，因此过程 $\pi^+ + p \rightarrow \Delta^{++}$（该过程 $I=3/2$）与过程 $\pi^- + p \rightarrow N$（$I=1/2$）的散射截面是不同的。费米的 π 子-核子散射实验同基于同位旋的理论预期一致，$\pi^+ + p \rightarrow \pi^+ + p$、$\pi^- + p \rightarrow \pi^0 + n$、$\pi^- + p \rightarrow \pi^- + p$三个过程的散射截面之比为9∶2∶1。

同位旋的地位自此确立了下来，无怪乎研究人员都将SU（2）对称性看做是与强相互作用相关的基本对称性候选方案之一。

但这种方案却没有成功。在试图建立一套基于SU（2）群的理论过程中，会出现一些巨大的数值。1954年，杨振宁和罗伯特·米尔斯（Robert Mills，图4.13）在美国布鲁克海文国家实验室提出，内在的SU（2）对称性应当在时空内的每一个点上单独成立，因此

	π^+	p	$\longrightarrow$	$\Delta^{++}(1232)$	
I	+1	$+1/2$		$+3/2$	
I_3	+1	$+1/2$		$+3/2$	
	π^-	p	$\longrightarrow$	$N^0(1520)$	$\Delta^0(1232)$
I	+1	$+1/2$		$1/2$	$3/2$
				或	
I_3	−1	$+1/2$		$-1/2$	$-1/2$

图 4.12 同位旋结合规律是一种向量叠加规律。只有第三方向分量 I_3 可以通过简单的方式叠加。同位旋预言，对于像 π^- p 这样的反应，不同的同位旋反应倒会产生不同的共振态

它成为了一种局域对称性。他们写道：

> “一旦某人在时空内的一个点上选择了什么叫质子、什么叫中子，那么以后这个人在时空的其他点上就不能随意地进行选择了。这看起来同作为通常物理基础的局域场概念不太一致。”

他们参考的原型是 QED 理论，在 QED 理论中要将全局 U（1）对称性强制变成局域对称性，从而引入了传递相互作用的规范场。在 QED 理论的情况下，规范场是光子。杨振宁和米尔斯希望在 SU（2）场的基础上进行同样的工作。他们最终没有建

图 4.13　杨振宁和罗伯特·米尔斯于 1999 年拍摄于纽约州立大学石溪市分校的一次研讨会上，这次会议后杨振宁宣布退休。杨振宁和米尔斯在 1954 年提出的理论框架成为了电弱理论和 QCD 理论的基础。图片来源：杨振宁

立起一套可行的强相互作用理论。但是他们的工作搞清楚了规范理论应该如何建立在非阿贝尔群（而不是以前所接触的阿贝尔群）的基础之上，因此产生了所谓的杨-米尔斯理论。其间的差别听起来好像只是一些很小的技术步骤，但由此引发的影响却是十分巨大的。实际上，电弱相互作用理论和 QCD 理论都是杨-米尔斯理论。杨振宁和米尔斯之所以没有发现这两套理论中的任何一套，主要是因为他们没有找到正确的对称性作为基础。

1956 年，普林斯顿高等研究院的日本理论物理学家内山菱友（Ryoyu Utiyama）对他们的工作进行了总结，以现在的观点来看，内山菱友的工作很明显是在他们理论的基础上，很有逻辑性

地一步步向现代场论前进。但在那时，他们的工作引起了一些争议。争议是由一些讨厌的问题所引发的，那就是规范玻色子质量的问题，这些争论几乎令整个思想毁于一旦。泡利在这一点上对杨振宁和米尔斯提出了质疑，他自己写出了类似杨-米尔斯理论的结果，但是没有发表。剑桥的研究生罗纳德·肖（Ronald Shaw）也得出了相似的结果，也没有发表。但杨振宁和米尔斯将结果发表了，他们的名字也因此同支撑QCD理论框架的基础理论联系在了一起。

家族壮大

核子、π子以及Δ共振态都是同位旋多重态，粒子家族通过对称性联系了起来——在以上的例子中，相关的对称性是同位旋对称性。这些家族的大小各不相同：核子形成的是双重态，π子是三重态，而Δ共振态是四重态。同位旋SU（2）对称性串起了一个家庭内的各种粒子，与各家庭成员数量的不同对应，相应的群表示也应当具有不同的大小。

这种思想又回到了我们之前所讲的那种桌面转动的情形，用2×2的矩阵很容易把它表示出来；如果转动还涉及桌面所在房间的墙壁，而不仅仅是桌面自身的边缘，那么就要用3×3的矩阵来表示。对于质子和中子来说，SU（2）群是通过2×2矩阵——这种群的最小表示实现的。SU（2）群的3×3表示将三个可能的π子联系起来。而4×4矩阵表示包括了所有四种Δ共振态。

这里有一点需要注意，将一种对称性应用于大小不同的粒子家族，对应着同一个群不同大小的表示，在以上情况中，群是指SU（2）群。SU（3）群和SU（4）群没有在这里出现，尽管它们的最小表示分别为3×3矩阵和4×4矩阵——这是因为它们是描述不同对称性和不同守恒定律的不同群。

找到这些更大表示的快捷方法就是先算出该粒子家族有多大。

这里用到的基本方法就是从较小的对称多重态出发，组合出更大的对称多重态。物理学家和数学家们为此推导出了一些定则，将以上所举的同位旋双重态结合的例子（例如两个核子结合）进行了系统整理。回想一下我们所得到的结果，有三个同位旋相同的粒子组成了一组，还有一个同位旋为 0 的单态，或者说是只有一个成员的粒子家族。这种关系可以写作 **2⊗2＝1⊕3**。其中的数字表示每个 SU（2）多重态“居住”了多少家庭成员。用黑体表示是提醒大家进行相加与相乘运算的是多重态。套着圆圈的特殊加号和乘号，表达的也是同样的意思。

实际上，表示矩阵本身看起来是什么样子是次要的。包含物理内容的是多重态，由于规则说明了怎样通过比较小的多重态形成比较大的多重态，物理学家们马上就知道了粒子家族是由对称性联系在一起的。这种将粒子“动物园”分门别类的方法十分有趣。

随着奇异性的引入，这些家族不断壮大。根据奇异性预言的 Σ^0 和 Ξ^0 粒子分别于 1955 年和 1959 年被发现，由此证明了重子数和奇异数都是与强相互作用相关的量。实际上，这两个量之和是除同位旋之外，强力的另一个守恒量。它也有自己的名字——重子数加上奇异数被称作超荷（hypercharge）。

用两个参量而不是一个参量，就意味着需要更大的对称群——比依赖于一个参数的 SU（2）群要大。完成这项任务的是出自同门的二参数群——SU（3）群。

人们之所以致力于研究 SU（3）群，部分源自于人们越来越期望理解已经观测到的粒子是怎样由更基本粒子组成的。1956 年，日本物理学家坂田昌一（Shoichi Sakata）假设质子、中子和 Λ 粒子是组成重子和介子的基本砖块；要引入奇异数机制，就必须出现 Λ 粒子。坂田昌一的目的是寻找隐藏在奇异性规则背后的意义，他的灵感来自于 25 年之前人们用中子解决了许多与原子

核有关的谜团，他推断 Λ 粒子就是人们尚未找到的那种粒子组分，就像中子曾经是人们尚未发现的原子核组分一样。

以今天的标准来衡量，认为 π 子是由核子和反核子结合而成的观点十分奇怪。但坂田昌一和其他一些人一样，都希望找到一些更基本的组分，从而解释“基本”粒子数量的激增。他们进行了一系列工作，后来的发展证明，他们的工作富有成效。

但是他们的努力并没有立竿见影。坂田昌一的模型一开始就搁浅了，这主要是由于没有找到这三种粒子怎样结合到一起的机制。1959 年，三位同样来自日本的物理学家取得了进展，他们三人分别是广岛的池田峰夫（Mineo Ikeda）和小川修二（Shuzo Ogawa）以及和坂田昌一一样来自名古屋的大贯义郎（Yoshio Ohnuki）。他们认真考虑了质子和中子是同一基本实体的两个不同态的思想，与之类比，提出了质子、中子和 Λ 粒子是同一物体的三个态，换句话说是一个三重态的三个成员。这三种粒子在某种意义上是相同的，至少在涉及强相互作用的情况下相同。所以，与同位旋 SU（2）群联系着两种核子一样，他们引入了 SU（3）群来描述 p、n 和 Λ 三重态之间更大的对称性。这三种粒子以及它们的反粒子便成了组成物质的基本“砖块”，像介子这样的粒子就是由它们“堆砌”而成的。

将 SU（2）群简单地推广为 SU（3）群，会造成两个之前没有预料到的变化。对于较小的群，其最小表示只有一种选择。但数学家们发现 SU（3）群的最小表示有两种，它们都用 3×3 矩阵表示，并且都同样有效。这两种表示存在于两个平行的数学世界里（可以用照相的例子作类比。相片是真实场景的一种反映，底片是真实场景的另一种反映，两种反映方式之间的联系并不是那样显而易见。）。因为它们都是 SU（3）群的最小表示，因此有两个多重态与之相关；由于它们都是 3×3 的表示，因此这两个多重态都是三重态。SU（3）的两个三重态之中的一个被写作

“3”，数字 3 说明 SU（3）三重态中存在着 3 种粒子；另一种三重态用星号标记，记作“3*”。池田峰夫和他的同事充分利用了这两种三重态的存在。由于粒子和反粒子也是不同的，一种存在于世界中，一种存在于反世界中，因此他们将一种三重态“3”确定为粒子，而将另一种三重态“3*”确定为反粒子。

还有另外一个问题。并不是所有的群都有任意大小的表示。例如 SU（2）群，从 2×2 矩阵开始，所有其他大小（或叫维度）的表示都是允许的。相应地，理论上也存在与SU（2)对称性有关的任意数量的粒子多重态。但 SU（3）群却不是这样，其可能的多重态只能取 3、6、8、10、15 等值。序列中有间隔是很正常的事情，只有 SU（2）群是一种特殊的情况。以上序列中的那些间隔表示，不存在通过 SU（3）对称性相联系的四重粒子家族，同样也不会存在五重粒子家族。

正如以前提到的，多重态结合规律会揭示更大的多重态。最令人感兴趣的一种结合是“3”态和“3*”态结合得到一个单态和一个八重态（octet）。用多重态的语言来表示，该结合可记作 **3⊗3* ＝1⊕8**。池田峰夫和他的同事将其作为坂田模型中一种三重态结合的方式引入，这迫使他们为 SU（3）群中的“8”指定了 8 种介子。这是他们仍存留至今的几种想法之一，但他们的模型却被逐渐遗忘。

将发生强相互作用的粒子指定给 SU（3）群的各种多重态，是一种很有成效的推理方式。至少伦敦皇家学院的一位不太年轻的研究生尤瓦尔·尼曼（Yuval Ne'eman）是这么想的。尼曼接触物理学很晚，这是因为之前他一直致力于建立特拉维夫地下犹太自卫队的工作。后来他参加了 1948 年的以色列独立战争。当时，他马上就会成为一名以色列内阁大臣，他在临时担当以色列驻英军士专员的同时，在阿卜杜斯·萨拉姆的指导下进行物理方面的研究。他研究的课题是在群论的帮助下，将粒子动物园的成

员进行分类。正是这项工作，使得尼曼和现在工作于加州理工学院的盖尔曼各自独立发现了“八正法”。

家庭的启发

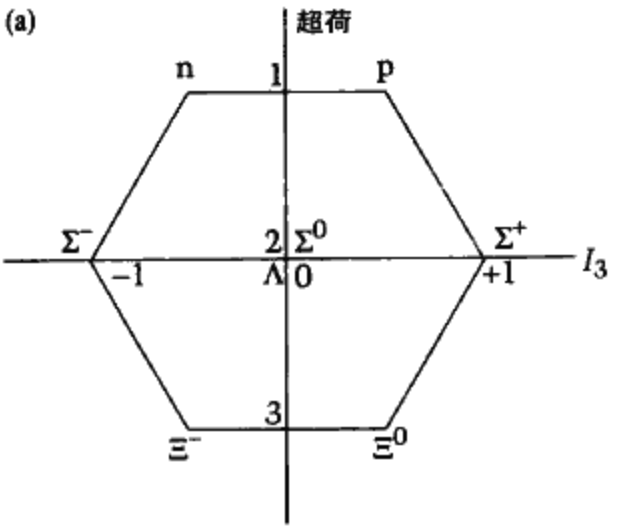
一张以同位旋和超荷为坐标轴，标出粒子的图表，会给一双双期望的眼睛带来一些启迪（图 4.14）。

第一幅图是八个重子的组成，第二幅是 π 子和 K 介子的组成。盖尔曼于 1961 年在“CTSL－20”论文中发表了这两幅图，论文发表在加州理工学院的内部期刊上（CTSL 代表 Caltech Synchrotron Lab——加州理工学院同步加速器实验室），但从未发表于任何其他的科学期刊。盖尔曼正是在这里为全世界引入了“八正法”这种表达方式。

盖尔曼和尼曼将这八个重子看做是质量基本相同的粒子，并假设它们都是某个八口之家（八重家庭）的成员。同样，π 子和 K 介子是另一个八口之家的成员。他们选出联系各个家庭成员的对称性是 SU（3）群，因为它是 SU（2）对称群的“最简单推广”，它的性质与其原型 SU（2）群最为相似。八重家庭最主要的性质反映在“八正法”的八上，这个概念同佛教中用来启迪人们思想的“八正道”有异曲同工之妙。

根据 $\mathbf{3}\otimes\mathbf{3}^{*}=\mathbf{1}\oplus\mathbf{8}$ 这种思想，人们有可能从更小的多重态构建八重家庭。因为其中存在 3^{*}，盖尔曼推测 n、p、Σ 以及 Ξ 一定包含反粒子。尽管最终证明这种想法是不正确的，但八正法的成功是显而易见的。尤其是在 π 子和 K 介子家族的引导下，盖尔曼和尼曼推测存在 η 介子，之后在同一年，就有人发表了首次观测到 η 介子的证据。他们还预言存在一个全新的粒子家族，该预言稍后又被证实了；他们还推断，在同一个八重家庭中的粒子应该具有某些共同的性质，人们后来发现这也是正确的。

盖尔曼和尼曼都在八正法的基础上，利用杨-米尔斯理论来



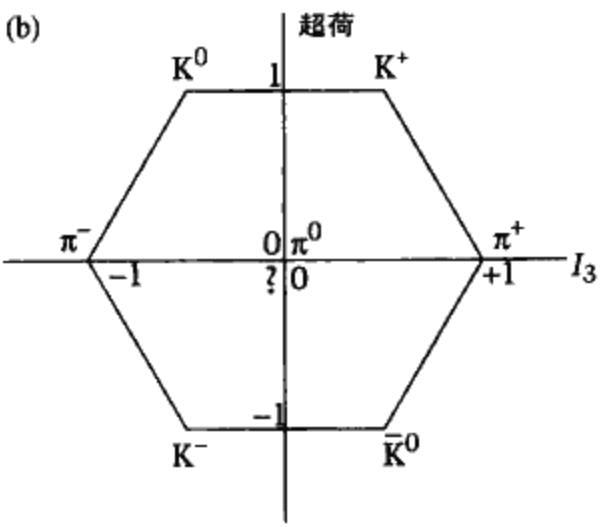


图 4.14　在 1961 年前后，人们用同位旋和超荷标记的已知重子（图 a）和介子（图 b）。纵轴上的数字代表超荷，在介子图中，由于重子数为零，超荷就等于奇异数。横轴表示同位旋的第三方向分量。介子图中出现的“?”，就是通过八正法预言存在的 η 介子。它在同一年被发现

考虑该问题，但他们两个都过于超前了——八正法没有提供一套有关强相互作用的完整理论。实际上，它暂时还仅仅是几种可选择方案之一。但 1962 年，人们根据八正法推断存在一种新的粒

子——Ω^-粒子，因此作为能理清粒子动物园关系的一种方法，八正法的地位确立了起来。此时，盖尔曼和尼曼指出Δ（1232）、Σ（1385）和Ξ（1530）共振态都应该属于某个十重家庭，以同位旋为横轴、超荷为纵轴将这些点画出来，会得到一个缺一个角的三角形（图4.15）。注意Σ（1385）和Ξ（1530）共振态要比出现在八重子图（图4.14）中的Σ和Ξ重一些。Ω^-粒子就是那个失去的角，因此它两个标签的值分别是$I=0$和$S=-3$（超荷是B与S的和，因此超荷为-2的一个重子$B=1$，这就表示奇异数$S=-3$）。他们还推测Ω^-粒子的自旋值为3/2，并且能够通过弱相

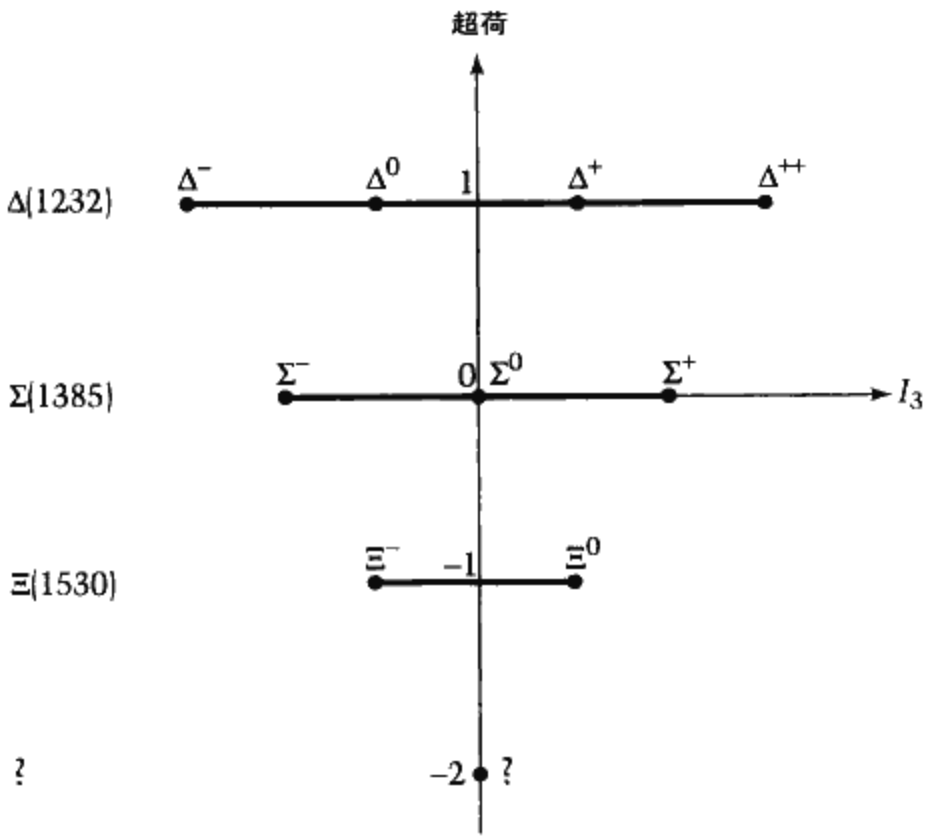


图4.15 顶点缺少一个粒子的重子十重态。那个空白应该被Ω^-粒子填充。Σ（1385）和Ξ（1530）都是共振态，它们分别同出现在重子八重态中质量较轻的Σ粒子和Ξ粒子相关

互作用发生衰变，因此它会比家族中的其他成员寿命更长。盖尔曼和尼曼通过研究发现，在多重态中粒子的质量实际上并不相同，而是存在系统上的差别，他们甚至能够估计出 Ω^- 粒子的质量大约为 1680 兆电子伏。两年后，人们在布鲁克海文国家实验室发现了 Ω^- 粒子以及预料之中的弱衰变现象——根据最新的实验结果，它的质量是 1672.4 兆电子伏（图 4.16）。

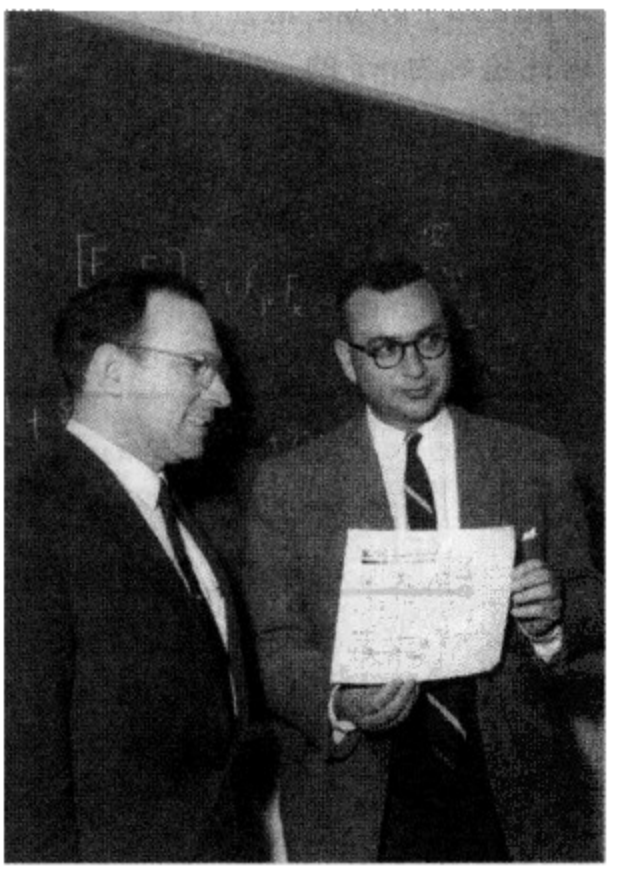


图 4.16　尤瓦尔·尼曼（左）和默里·盖尔曼第一次揭示 Ω^- 粒子的著名照片，拍摄于 1964 年的加州理工学院。Ω^- 粒子的发现为八正法提供了确凿的证据。图片来源：加州理工学院图片档案室

叫来马克的夸克

“酉群［SU（3）］的数学表述是通过讨论三个假想的‘轻子’

实现的……这与真实的轻子或许有关，也或许没有关系，”盖尔曼在他 1961 年发表的一篇有关八正法的报告中写道。盖尔曼认为，讨论 SU（3）对称性以及用三种假想粒子讨论其作用，“对物理学思想的发展很有帮助”。

他将他提出的三种假想粒子分别叫做电子、μ 子和中微子，为了简单起见，他假定它们的质量都是 0，它们可能与具有相同名字的粒子有一定联系，但也可能没有。这样，这三种东西就变成了 SU（3）三重态的成员。随后，盖尔曼通过给他的假想费米子设定一定的质量，解释了重子和介子的质量差异。

存在三种具有 SU（3）对称性费米子的假想概念非常重要，人们可以用它来理解同样具有 SU（3）对称性的八重子态。因为八正法隐藏了一个谜团。具有重要意义的多重态看起来像是十个一组的十重态——为什么基本的三重态在自然界中看不到呢？这可以同作为原型的同位旋作对比，尽管同位旋双重态同更大的多重态有很多联系，但双重态实际上只是中子和质子。简言之，在 SU（2）群中由中子和质子所担当的角色，在 SU（3）群中由谁来扮演呢？

在一个简单的数学结论中隐藏着一条线索。三套 SU（3）三重态会产生十重态，并且还带来许多其他的东西。实际上，完整的过程应该是 $\mathbf{3}\otimes\mathbf{3}\otimes\mathbf{3}=\mathbf{1}\oplus\mathbf{8}\oplus\mathbf{8}\oplus\mathbf{10}$。1964 年，盖尔曼提出了三种基本实体，取代了陈旧的假想轻子，它们就是出现在乘法当中的 3。这样，因为有三套这样的三重态，因此也需要三种这样的实体来构建重子的十重态。他最初把这种实体称作“kwork”，后来改成了“quark”（夸克），取自于詹姆斯·乔伊斯（James Joyce）的书《芬尼根守灵夜》(*Finnegan's Wake*)①。盖尔曼论文的参考文献中列出了乔伊斯的书，这种在物理杂志中出现经典文

① 译者注：原文为 Three quarks for muster Mark。

学的情况即使不是仅有的一次，也是很罕见的。实际上，夸克也是一种德国软奶酪的名字。

盖尔曼把他的三类（或者称作三种“味道”）夸克分别称作上夸克、下夸克和奇异夸克，或者是 u 夸克、d 夸克和 s 夸克，这些名字一直沿用到今天。在这幅图像中，重子由三个夸克（qqq）组成。根据量子场论，它们像电子和其他费米子一样，存在相应的反粒子，即夸克一定同时伴随着反夸克存在。反夸克用 $\bar{q}$ 表示，或者可以将三种味道夸克的反夸克分别表示为 $\bar{u}$、$\bar{d}$ 和 $\bar{s}$。根据这种方案，介子是由夸克和反夸克对组成的，表示为 $q\bar{q}$。将三种夸克同 SU（3）群中的“3”联系起来，反夸克就对应着“3^*”，结合成重子与介子的相应多重态分别是 **3×3×3** 和 **3×3***。在该方案中，质子由三个夸克组成，其中两个上夸克、一个下夸克，记作 uud；而中子是 udd，Σ^+ 粒子的夸克组成方式是 suu。图 4.17 说明了八正法中重子和介子十重态的夸克分配情况，以及重子所组成的十口之家（或称十重态）。这种粒子的“夸克模型”和旧“坂田模型”之间存在着惊人的相似之处——坂田昌一只是用错了那三种构筑物质世界的砖块，并且通过错误的方式把它们结合到了一起。

但得到夸克模型却付出了很高的代价。盖尔曼曾一度将一种他不太满意的性质，当作理论发展的障碍。为了让该理论行得通，夸克必须带有分数电荷。u 夸克所带电荷为单位电荷的 2/3，s 夸克和 d 夸克都是－1/3，它们所带电荷都是电子电荷的分数倍。这种分数电荷情况同人们以往得到的经验——所有电荷都是基本电荷的整数倍——相抵触。具有 1 倍、2 倍、3 倍以及更多倍于电子电荷的情况都能令人满意，但是，如果存在 1/3 或 2/3 倍电荷就要推翻基本定律。令人头痛的还不只这些，由于重子是由三个夸克组成的，还必须要求每个夸克的重子数都是 1/3。这些令人吃惊的结论就是夸克模型所遇到的障碍。

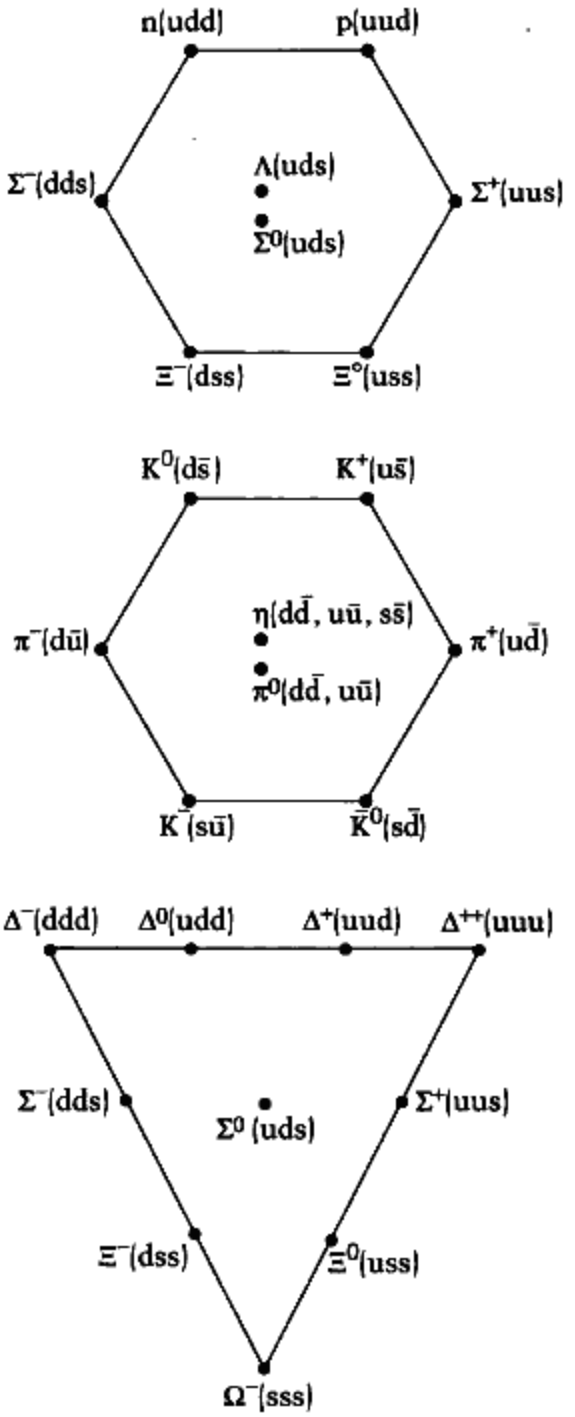


图 4.17 重子八重态、轻介子八重态和重子十重态的夸克模型表示

但夸克的其他性质都比较令人满意。夸克是自旋为 1/2 的费米子，在这方面它满足八正法中要用到的三种假想费米子条件。u 夸克和 d 夸克形成了一个同位旋双重态，其同位旋第三分量分别是 $I_3=+1/2$ 和 $I_3=-1/2$，这也可以解释中子和质子的同位旋性质。s 夸克没有同位旋，但它的奇异数是－1（表 4.2）。

表 4.2　　夸克和反夸克的性质

夸克/反夸克	电荷	同位旋	奇异数	重子数	自　旋
u	+2/3	+1/2	0	+1/3	1/2
d	−1/3	−1/2	0	+1/3	1/2
s	−1/3	0	−1	+1/3	1/2
$\bar{u}$	−2/3	−1/2	0	−1/3	1/2
$\bar{d}$	+1/3	+1/2	0	−1/3	1/2
$\bar{s}$	+1/3	0	+1	−1/3	1/2

八正法可以理解为，质子和它所在家族的成员都是由三部分组成的。其他人的脑海中也存在过这种思想。早在 1962 年，尼曼和他的同事海姆·戈德堡（Haim Goldberg）就意识到十重态与三个三重态有关，这说明质子存在三个组分。但他们也意识到，这就意味着要出现分数量子数，这种想法看起来很可笑，因此他们认为最好只将这些组分当作一种计算工具。

但毕业于加州理工学院的乔治·茨威格（George Zweig）并没有躲在理论的障碍背后。茨威格在上研究生的时候，听盖尔曼讲过他的八正法。后来，茨威格在 CERN 独立提出了自己的夸克理论，他将夸克称为埃斯（ace），并将其视作非常真实的质子组分。茨威格的理论从未在物理期刊上发表过，只是在 1964 年初发表了两篇 CERN 内部文章。这对夸克模型来讲是一种损失，也将他的贡献汇入量子理论的时间推迟了很多年。

盖尔曼自己在那些受争议的问题上宣称，夸克是不能被单独看到的，只能作为像中子和质子这样性质容易理解的粒子的一部

分。以这种方式来看，分数电荷和分数重子数就显得不那么“真实”了。但是可怜的夸克到底应该被置于一个什么样的位置呢？它是一种实验者能够找到的粒子呢，还是用来解释群论问题的一种神秘而有效的数学工具呢？

盖尔曼在1964年写道：

> “如果夸克是一种具有有限质量的物理上的粒子（而不是纯粹的可能有无穷质量的数学概念），推测它们的行为方式是一种有趣的事情。”

在电荷和重子数的确守恒的基础上，盖尔曼推测至少有一种夸克是稳定的。他提出，地球表面上可能会存在一小层由宇宙射线引起的、稳定的夸克物质污染物。最后，他写道：“对稳定夸克的寻找……高能加速器会帮助我们证明不存在真实的夸克。”根据盖尔曼的思想，“数学上的”夸克是一些不能单独出现、也不能被单独看到的东西；“真实”夸克是能够独自产生和观测到的。这是一个奇妙的想法，这种想法在他头脑里萦绕了很久。他在1992年写道：

> “我所谓‘数学上的’夸克，就是那些现在普遍认为真正存在的，并且由QCD理论预言了的东西。然而直到现在，许多权威人士仍坚持认为：我当初提出夸克的时候，很可能只是‘数学上的’，而不可能是‘真实的’，我的意思应该是它们并不存在。我当然不是那个意思。”

抛开语义学上的纠缠，夸克确实是存在的，问题只是在于我们是否能以令人满意的方式观测到孤立的夸克。如果存在自由夸克，实验物理学家们所面临的挑战就是发现它们。

夸克现身

有许多零散的证据能够证明夸克是真实存在的，它们是质子与中子这样的粒子的组成部分，唯一棘手的问题就是没有人“亲眼”看到过夸克。一种能够证明三夸克重子构成的精妙实验，利用了夸克具有的一种像小磁铁的性质。

1995 年，费米实验室的实验物理学家们经过多年努力，终于完成了精确测量由 u 夸克、d 夸克和 s 夸克组成的所有重子磁矩的工作（图 4.18）。他们的数值表中，列出了中子、Λ 粒子和 Ξ^0 粒子磁矩的测量值。如果这三种中性粒子是基本的点状物体的话，它们的磁矩应该是零，因为基本粒子的磁矩只与旋转的电荷有关。但是费米实验室的测量揭示了这三种粒子的磁矩值分别是 −1.913、−0.613 和 −1.253（磁矩单位称作核磁子，是根据相对论量子力学算出的点状质子所具有的磁矩值）。这些值不仅同零的差别非常大，而且互相之间也有差别。这说明这些粒子都是由带电的组分组成的，其总电荷为零，但由于组合方式不同，所得的磁矩也是非零的不同值。

实际上，解决质子和中子的磁矩问题是夸克模型的一次胜利。简单地假设每个夸克都是点状、旋转的带电物体，根据自旋为 1/2 的物体结合的标准规律，可以把组成粒子的每个夸克的磁矩叠加起来。质子磁矩的计算值同中子磁矩的计算值的比值为 −2/3，同实验测量值 −0.685 吻合得很好。对于由 u 夸克、d 夸克、s 夸克组成的 Λ 粒子，夸克模型指出，其磁矩只与奇异夸克的磁矩有关，同时 Ω^- 粒子的磁矩正好是奇异夸克磁矩的 3 倍。因此夸克模型预言，Λ 粒子同 Ω^- 粒子磁矩的比例应该是 +0.33；根据 1995 年公布的 Ω^- 粒子磁矩的精确测量值，实验给出的值是 0.304。两者吻合得也相当好。运用质子、中子和 Λ 粒子的磁矩测量值，很容易算出这三种夸克各自的磁矩来；u 夸克、d 夸克、

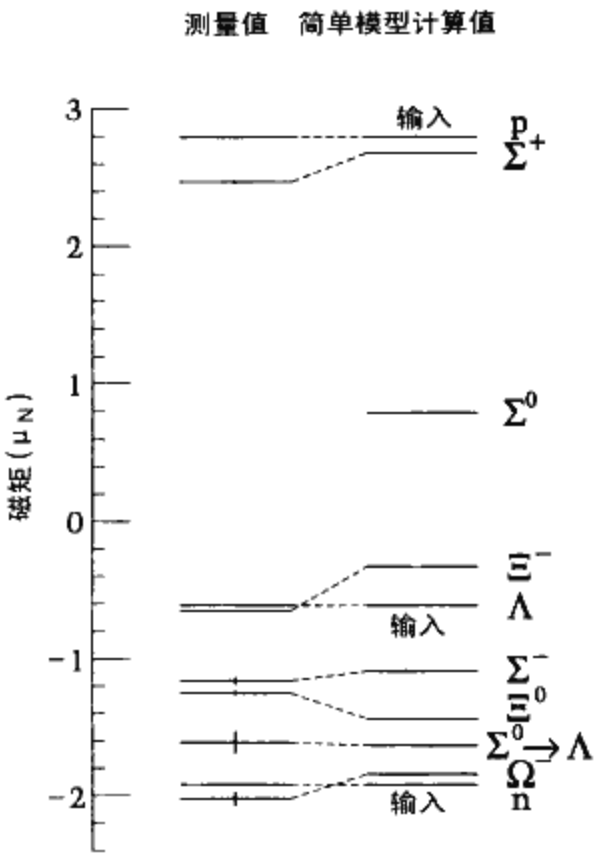


图 4.18 重子磁矩的测量值同简单夸克模型的预言值相对比。图片来源：粒子数据小组

s 夸克的磁矩分别是 +1.852、−0.972 和 −0.613。然后，再用夸克模型去计算其他重子的磁矩也同样很容易。总的来说，除了 Ξ^- 粒子之外，计算值和测量值在 10% 的误差范围内都吻合得很好，Ξ^- 粒子的误差则在 30% 之内。对于这样简单的计算方法，这种吻合程度已经十分令人满意。

除了这些常数之外，点状旋转粒子的磁矩还依赖于它们的电荷和质量。这就说明，如果重子内部的夸克真的是旋转的点状粒子，实验者们就可以认真考虑列出的每个夸克的磁矩，从而得出三种夸克的质量。以兆电子伏为单位，它们的质量分别是：u 夸克 338，d 夸克 332，s 夸克 510。这些质量都是“有效质量”，它

表示无论夸克是否能自由存在，如果由夸克们组成的重子中真的存在夸克，那些夸克所呈现出的质量就是有效质量，即把它们当作自由夸克的质量。

所有这一切，使得重子磁矩实验具有很大价值。中子、Λ粒子和Ξ°粒子这些中性粒子具有有限的磁矩，证明了它们都由带电的组分构成。利用简单夸克模型进行计算的基础，是旋转的带电夸克和简洁的对称性原理，它不涉及整个系统怎样结合在一起的复杂理论，但最终却能够给出与实验十分吻合的结果。磁矩研究不仅证实了夸克模型的可靠性，而且还顺便给出了夸克本身有效质量的估计值。

彩色的夸克和白色的光子

夸克作为质子、中子以及其他组合体的下层结构，前景看起来一片光明，但是在20世纪60年代中期，理论物理学家手头上还有一个鲜活的问题需要解决。夸克模型仍然存在一个小的技术缺陷，但它的意义却十分重大。这是一个盖尔曼已经知道，并且为之忧心忡忡的问题。这个问题就是：粒子的统计性质不对。

夸克必须是费米子。作为费米子，它们就要遵从不相容原理；只有对于费米子，讨论由确定数目的夸克组成质子和中子这种问题才有意义。如果夸克是玻色子，不相容原理就不起作用，任意数量的夸克都可以组成一个复合粒子。比如，因为光子是玻色子，组成激光的光子不能被一个个辨认出来，对于玻色夸克，它们互相之间也不能区分，夸克模型就失效了。

此外，如果想要满足由夸克组成物体的自旋性质，夸克的自旋也必须是1/2。由包括质子和中子在内的重子组成的八口之家，其成员的自旋值都是1/2，而重子十重态的自旋都是3/2，介子的自旋是0或1，它们都可以看做由相同自旋值的基本砖块——夸

克所组成的。实际上，夸克自旋是夸克模型的根本内容。这也体现在粒子统计问题中。

该问题在重子中体现得十分明显，最清楚的例子是自旋为 3/2 的 Ω^- 粒子——它是十重态中的一种粒子。Ω^- 粒子是由三个奇异夸克（sss）所组成的，由于组成它的三个夸克是相同的，因此其结构是对称的，如果交换任意两个夸克将不会发生任何改变。要使总自旋值为 3/2，所有三个夸克都要自旋向上（图 4.19），这同样是一种对称结构。实际情况就是这样，夸克的味道和夸克自旋的结合满足对称性。但怎么会出现这种情况呢？Ω^- 粒子本身是一种费米子——其自旋为 3/2，因此它必须满足整体的反对称性。

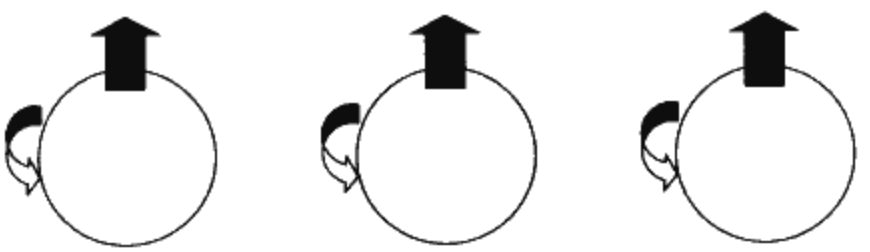


图 4.19　在 Ω^- 粒子中，三个自旋方向相同的奇异夸克不可分辨

Ω^- 粒子是一种特殊的情况，只有在这种情况下才容易看出问题。自旋为 3/2 的 Δ^{++} 粒子也存在同样的问题，它是由三个上夸克（uuu）所组成的；具有同样问题的还有由三个下夸克（ddd）组成的 Δ^- 粒子。这两种粒子都同 Ω^- 粒子一样，具有明显对称的夸克和自旋结构，因而给出错误的对称性。实际上，在重子十重态中，所有 10 种重子中的夸克结合对于交换两个夸克的操作都是对称的，即使那些由上夸克、下夸克和奇异夸克混合构成的粒子也是这样。例如 Σ^0 粒子是由三种夸克组成的，见图 4.17，这种表示实际上是 uds＋dsu＋sud＋sdu＋dus＋usd 与一些数值常数相乘的缩写。由于家族中所有 10 位家庭成员的自旋都

是 3/2，它们的自旋结合都是对称的，因此整个家族的整体对称性都是“错误的”。

如果不计自旋（以及夸克之间的互相绕转），重子八重态总体上不全是对称的，也不全是反对称的。值得注意的是，当自旋卷入这幅图像中时，八重态之间的“混合”对称性，精确地对应于与之相似的三夸克自旋混合对称性。结果会导致整体上的对称性结合：重子八重态的所有八位成员——包括质子和中子，在夸克的味道和自旋上都是对称的。这是非常巧合的事情，因为每种夸克都具有三种可能味道中的一种，并且要么自旋向上、要么自旋向下；它们能组合出 91 种“简单”重子，其自旋值为 1/2 或 3/2。然而在这些候选者中，只有那些明显具有错误整体对称性的粒子才被挑选出来，用来组成自然界中的事物。

多彩的解决方案

普林斯顿高等研究院的奥斯卡·格林伯格（Oscar Greenberg）设计了一套解决方案。他的想法是，坚持三个一组的夸克其整体行为类似于费米子。换句话说，根据格林伯格的思想，夸克要遵从一套经过修正的费米子统计规则，三个夸克组成的夸克组，整体上满足正常的费米统计，即使夸克组中的夸克具有对称的结构也不例外。用这种方法，他的方案能自动保证重子满足正确的整体统计性质。其代价就是要给每个夸克添加一种新的标签，该标签的值可以取三个可能值中的一种（图 4.20）。

美国纽约州锡拉丘兹大学的韩武荣（Moo-Young Han）和芝加哥大学的南部阳一郎（Nambu）也设计出了一套类似的解决方案，他们的方法中也包含一种具有三个值的新标签，他们同格林伯格一样在 1964 年发表了他们的想法。韩武荣和南部阳一郎给每个夸克附加了一种可取三个值的标签，并假设“普通”重子——重子八重态和重子十重态中的成员所具有整体反对称性是

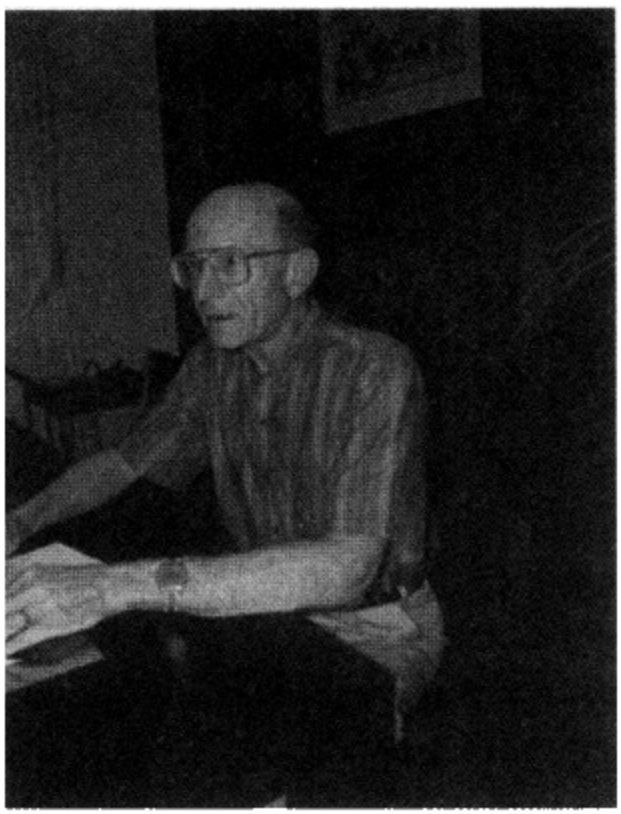


图 4.20　奥斯卡·“傻瓜”·格林伯格，2003 年摄于马里兰州的家中。格林伯格引入了色属性，将其作为解决夸克态统计问题的一种方法。最终证明，色荷的作用是 QCD 理论的核心内容，就如同电荷是 QED 理论的核心一样。但是通往优秀思想的道路并不总是一马平川的。格林伯格回忆了当时的情景，他将论文拿给时任普林斯顿高等研究院院长的罗伯特·奥本海默看，奥本海默说：“你的论文很优美，但是我不会相信论文中的任何一个字。”图片来源：奥斯卡·格林伯格

由这种新标签引起的。由于该标签能取三种可能的值，他们进一步假设，这种标签像夸克的三种味道一样，也应该满足 SU（3）对称性：他们在夸克图像中引入了另一种 SU（3）对称性。

韩武荣和南部阳一郎不喜欢夸克具有分数电荷的想法，因此他们尝试在模型中确保夸克电荷为整数。我们将韩武荣和南部阳一郎的观点修正一下，使之符合今天所流行的观点，即夸克所带的电荷确实

是分数倍的，他们的模型就和格林伯格的模型完全等价。附加给夸克的这种新标签现在有一个名字，叫做颜色。

颜色是一种具有三个可能取值的新量子数。你会把一个新量子数叫做什么？你可以发明一个全新的词汇来表示它，或者从经典文学中抽象出一些没有意义的词（就像“夸克”的情况一样），或者就使用一个像“颜色”这样已有的词汇。使用“颜色”所带来的问题是，有些人可能会认为夸克真的具有可染色的性质，但实际不是这样。颜色仅仅是夸克具有的某种量子属性的名字，但将其名字定为“颜色”实际上十分方便。某个单独的夸克可能是红的、绿的或蓝的，对应于三个可能的颜色标签值，用这些颜色名来命名这三个值实际上只是为了方便。组成一个重子的三个夸克，其颜色标签各不相同，这就说明这三个夸克现在不再一样了，这就是用颜色解决重子统计问题的关键。例如组成 Ω^- 粒子的三个奇异夸克，一个是红色的，另一个是绿色的，第三个是蓝色的（图 4.21）。

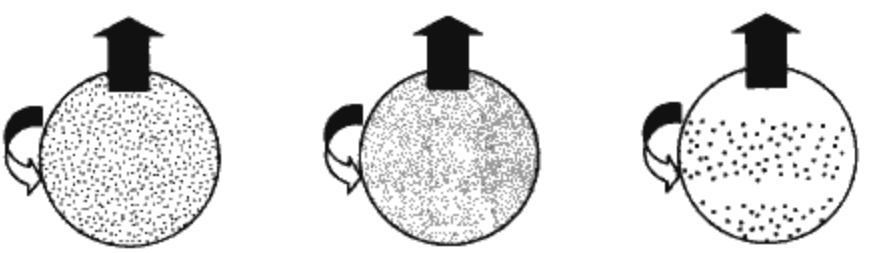


图 4.21　组成 Ω^- 粒子的三个自旋方向相同的夸克，现在可以通过它们之间颜色的不同加以分辨了。它们分别是红色、绿色和蓝色的，图中用不同的阴影进行区别

第二套 SU（3）群到底是什么样的？引入框架中的两套 SU（3）群真的有什么奇妙之处吗？引入第一套 SU（3）群是因为，在那些引出夸克模型的普通重子中，需要存在三种不同类型的夸克。需要两种味道来区别中子和质子，第三种味道引入新发现

的奇异性。这套与夸克的味道相关的SU（3）群有时记作$SU(3)_f$或SU（3）-flavor[1]来强调它的作用。尽管它被称作“f”标签——或者“味道”标签，但在数学上它仅仅是一种简单的SU（3）群。如果存在的夸克不止三类——现在就有这样的理论——表示味道的$SU(3)_f$群将被更大的SU群替代。另一方面，表示三种颜色的SU（3）群通常写成$SU(3)_c$，从而与表示味道的$SU(3)_f$相区别。之所以要引入它，是由于重子中存在三个（很可能相同）的夸克，因此需要有三个不同的标识来区别它们。

然而即使引入了颜色属性，还是有一些不太对的地方。格林伯格的方案看起来有点将问题复杂化了，它允许存在大量我们尚未观测到的粒子——在1964年的论文中，他列出了满足对称群要求的2600种重子。韩武荣和南部阳一郎同样也推出了许多未观测到的粒子，但他们的论文已经提供了能够进一步发展的理论雏形。他们认识到，有一种反对称的颜色单态，是解决普通重子对称性问题的关键——这个问题我们稍后再详细说明。但他们却没有对此采取任何措施，来限制具有非颜色单态结构重子的其他可能性。他们像格林伯格一样，允许存在真正可观测的夸克，这意味着实验能够观察和测量颜色属性。

经过大量实验的验证，颜色属性的地位逐渐稳固，但实验者们从未直接观测到颜色属性。这是因为大自然规定，能够被观测到的粒子整体上不应该带有颜色。根据这条关键原理，由于夸克带有颜色荷，因此夸克本身不能够被观测到。重子由三个夸克所组成，每个夸克都带有不同的颜色。像质子这样真实的可观测重子，是通过带颜色的夸克进行量子混合得出的，整体上看，它是一种“白色”的复合粒子，整体上没有颜色。

---

① 译者注：flavor是“味道”的意思，f是其缩写。下文的角标c是英文colour的缩写，colour是“颜色”的意思。

这种思想对介子也适用。可观测的介子是夸克和反夸克的量子混合体。与电子所带电荷为－1，而其反粒子——正电子带有值为＋1的“反电荷”类似，反夸克也同样带有“反颜色”，分别用$\bar{r}$、$\bar{g}$和$\bar{b}$来表示。介子中，夸克和反夸克所带的颜色和反颜色相互抵消，得到“无色的”可观测粒子。对于该量子数，颜色确实是一个有用而形象的名字。

这种思想是韩武荣和南部阳一郎提出的颜色单态原理的发展——真实的可观测粒子都是“颜色单态”。这种反对称单态的原型早已出现过，只不过那时是作为两个自旋（或同位旋）为1/2的粒子按｛A（上）B（下）－B（上）A（下）｝方式结合的形式出现的。颜色通过相应的方式进行特有的结合，也可以产生相似的单态——即一种“白色”的态。就像夸克的三种味道会形成一个SU（3）$_f$三重态（或3态）一样，三种颜色（分别标记为r、g和b）也会形成SU（3）$_c$三重态（或3态）。因为重子是由三个夸克组成的，在重子的总颜色上将会体现出三套颜色3态的特征。这三种3态结合为**3⊗3⊗3＝1⊕8⊕8⊕10**。其中的**1**就是最为重要的颜色反对称单态，它对应着真实的粒子。明确地说，重子的颜色反对称单态应该是｛rbg－brg＋bgr－gbr＋grb－rgb｝，这种方式将各种颜色正确地混合在一起，从而确保形成“白色”的粒子。其他方式的颜色组合都不会在自然界中出现。

由于介子是由夸克和反夸克组成的，因此情况有所不同。三种反颜色$\bar{r}$、$\bar{g}$和$\bar{b}$存在于SU（3）$_c$群的3*三重态当中。这些态将会和颜色三重态3态通过**3⊗3*＝1⊕8**结合。其中的**1**就是颜色单态，可确定表示为｛$r\bar{r}$＋$g\bar{g}$＋$b\bar{b}$｝，其中夸克所带的颜色和反夸克所带的反颜色都可以互相抵消。这种单态的结合方式具有颜色的对称性。而另一方面，其中的8态总体上具有净颜色。它扮演了一个特殊角色，这部分内容会在适当的地方详细说明。

出现颜色SU（3）$_c$群，是因为它能使三个夸克组成一个重

子。出现味道 SU（3）$_f$ 群是因为，在组成重子的过程中只有三种夸克供我们选择，至少最轻的那些重子是这样的。味道 SU（3）$_f$ 对称性表示，将质子和中子结合起来的强力，对夸克三种味道的混合（u↔d↔s）十分迟钝。这就是为什么具有（三种味道中）任何一种味道以及适当自旋方向的三个夸克，都能给出同一多重态（八重态或十重态）中的重子。实际上这种对称性只是近似的。在给定的多重态中会存在某种形式的质量差异，这说明“严格”的对称性存在某种“破缺”。颜色SU（3）对称性以及能观测到的粒子必须是颜色单态的要求意味着：对于任何重子，将夸克结合在一起的力，在夸克三种颜色任意混合（r↔g↔b）的情况下不发生改变。夸克的味道和夸克的颜色是不同的：一种味道的夸克都可能具有三种颜色中的任意一种，如表 4.3 所示。

表 4.3　u、d、s 三种味道的夸克中的每一种，都能具有 r、g、b 三种可能颜色的任意一种

味　道	颜　　色		
	$u_{红}$	$u_{绿}$	$u_{蓝}$
	$d_{红}$	$d_{绿}$	$d_{蓝}$
	$s_{红}$	$s_{绿}$	$s_{蓝}$

颜色的提出以及能观测到的粒子必须是颜色单态的规定不仅解决了粒子的统计问题，还揭示了许多其他的内容。例如在考虑颜色问题之后，像 qq、qq $\bar{q}$以及 qqqq 这些组合方式都是被禁止的，现实中也确实没有看到过这些粒子的存在。所以并不是所有可能的 SU（3）$_f$ 多重态都能形成粒子。例如味道多重态 6 就不存在粒子成员。它是从味道 **3**⊗**3** 结合得来的，对应着被禁止的 qq 态。

颜色荷

解决重子夸克模型的统计问题，仅仅是颜色属性作出的一部

分贡献，更精彩的内容马上就要出现了。夸克间奇妙的颜色对称性，还为一整套新的规范理论提供了基础，这种规范理论的基础是局域颜色对称性，理论中会有一系列带颜色的规范玻色子在带颜色的夸克之间传递相互作用力。这种想法来源于20世纪70年代早期，一个思想十分活跃的时期，它构成了一种全新的理论——一种被称作量子色动力学（QCD）的理论的基础。自然界四种基本力中的最后一种被征服了。

这种基本思想就是，重子中的每个夸克都具有三种可能颜色之一，两个夸克之间的强力源自于它们所带的颜色，这就同两个电子之间的作用力源于它们所带的电荷一样。电力是由光子传递的。与之类似，一种新的携带力的粒子——胶子（gluon）传递着与色荷有关的力。量子色动力学的名字最初是由盖尔曼起的，其中的“色”（chromo）是指颜色，它在理论中扮演着一种基本的角色，是夸克间作用力的起源。

在描述强力过程中必须包含某种胶子，这个思想已经存在了好多年。早在1962年，坂田昌一模型提出之前，盖尔曼在一个模型中就提到了强子组分之间的“强胶子耦合”。但是依照QCD理论，只有一种胶子是不够的，理论需要存在一套由八个电中性、携带颜色的胶子组成的胶子八重态。胶子自身能够携带颜色的事实，具有十分重大的意义。首先，它意味着裸胶子同裸夸克一样，不能够被看到。也就是说，由于夸克和胶子都带有色荷，根据能够实际观测到的粒子不应具有净色荷的原理，直接观测夸克和胶子都是不可能的。

它还表示，一个胶子能够感觉到另一个胶子的存在，因此胶子之间能够直接发生相互作用。这与光子完全相反，光子不带电荷，一个光子不能直接同另一个光子发生电相互作用。从最简单的角度来理解，光子之间不能发生相互作用。

胶子携带颜色，因此彼此之间能够直接发生相互作用，发生

该现象的源头可以一直追溯到 SU（3）群的非阿贝尔特性。从理论物理学家们的角度理解，光子彼此之间不发生直接的相互作用，是因为 QED 理论的对称群是阿贝尔群。这在研究群论的数学家们看来，都是很不起眼的细节问题，但它们最终给自然界中相互作用的基本性质造成了很大影响。

SU（3）群的其他基本性质，也直接转变成了 QCD 理论的其他重要特征。尤其是 SU（3）群依赖于 8 个独立参数的特征，在 QCD 理论中转变为 8 种守恒荷，每一个群参数对应着一种荷。

规定电子-正电子场满足局域 U（1）不变性，表示光子的自旋为 1。为了保证对称性，从一个点转移到另一个点的费米子场一定要连续变化。这种变化将通过光子实现。将规范理论应用到夸克颜色对称性上，并使对称性局域化，就会取得同样的效果，但所用的对称群更加复杂。规范理论会在 8 种胶子的基础上生成规范相互作用，每种胶子都带有一种守恒荷，并且，所有胶子都是自旋为 1 的无质量玻色子。在效果上，8 种胶子都像是带颜色的光子，但它们与夸克所带的颜色发生耦合，而与夸克所带的电荷不发生作用。

颜色混合与匹配

图 4.22 的图（a）用形象的方式表现了电子之间的相互作用，其中一个电子向另一个投掷光子，而光子被另一个电子所吸收。电子只能携带一种可能的电荷，其反粒子也相应地携带一种反电荷，光子根本不带电荷，所以在 QED 理论中出现的可能性相对有限。

QCD 理论的结构要复杂得多，它有 3 种颜色和 3 种反颜色，还要加上带颜色的传递粒子。同简单的 QED 理论相似，一个红色夸克能够投掷胶子而自身仍保持红色，见图 4.22 的图（b），图中的胶子用绕圈的螺旋线表示，而光子则用波浪线表示。但这

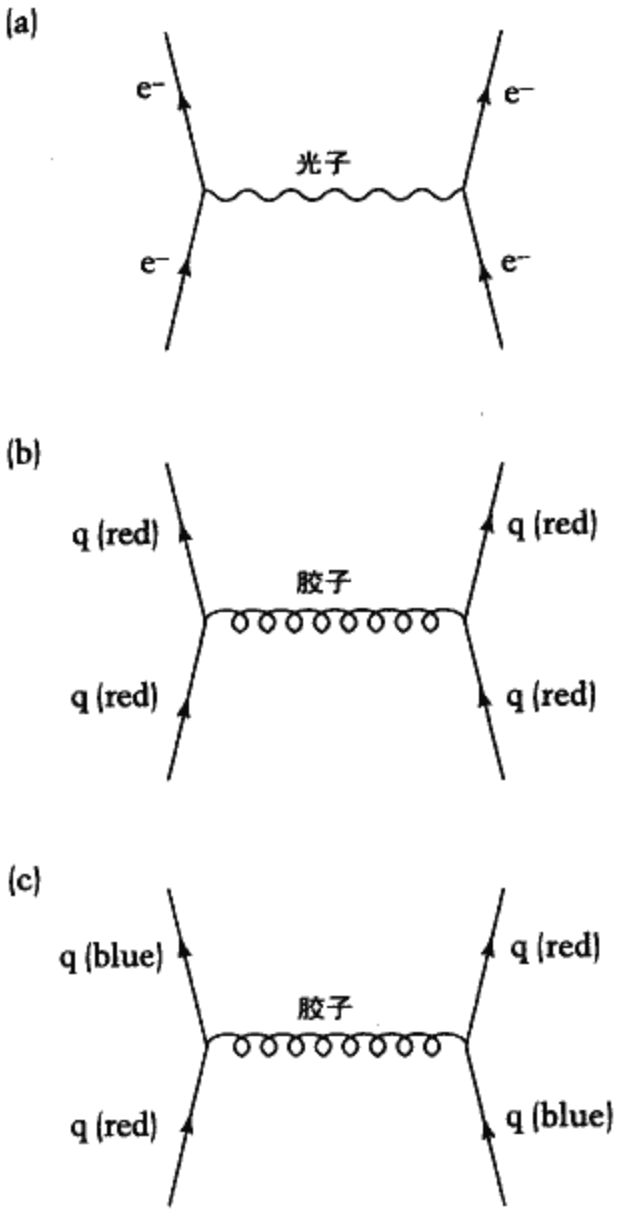


图 4.22 图（a）：通过交换光子能传递简单的 QED 相互作用。图（b）：在 QCD 理论中也存在交换胶子的等价过程。图（c）：这种过程还包含夸克颜色的交换

个过程也会引发奇异的变化。尤其是夸克的颜色会变为蓝色或绿色，“颜色变化”会通过胶子传递给吸收胶子的夸克，从而改变

接收夸克的颜色，见图 4.22 的图（c）。

要保证颜色守恒，可能出现的胶子颜色结构必须与通过胶子相互作用所引起的夸克颜色变化相对应。在夸克颜色发生相互作用的 9 种可行方案之中，有 6 种是发生颜色变化的：r→g，r→b 以及其他 4 种，3 种是颜色不发生改变的：r→r，g→g 和 b→b。在这 3 种颜色守恒的过程当中，只需要 2 种（任何 2 种）就足够了。第 3 种情况能通过交换颜色单态胶子实现。但这在 QCD 理论的规则下是不允许的，因为颜色单态的物体不会受到颜色力的作用。根据以上的分析，只剩下 9 减 1——8 种胶子颜色的方案。因此胶子应该是颜色八重态。在 $\mathbf{3}\otimes\mathbf{3}^{*}=\mathbf{1}\oplus\mathbf{8}$ 结合中就存在一个八重态。这就是我们之前提到过的携带颜色的八重态，它确实起到了特殊作用：它容纳了胶子。

完整的胶子颜色系列共有 8 种组合，分别为：$r\bar{g}$，$b\bar{g}$，$g\bar{b}$，$r\bar{b}$，$b\bar{r}$，$g\bar{r}$，$\frac{1}{\sqrt{6}}(r\bar{r}+g\bar{g}-2b\bar{b})$，$\frac{1}{\sqrt{2}}(r\bar{r}-g\bar{g})$。前 6 个很容易理解，它们来自于图 4.23 的图（a）所示的夸克-夸克相互作用，其中“逆流前进”的颜色是用相应的反颜色来表示的。图 4.23 的图（b）说明了夸克-反夸克相互作用中的颜色流动。最后 2 种胶子颜色组合，尽管缺乏物理直观性，但仍可以很明显地看出它们确实携带颜色。

并不是所有能想到的夸克和胶子的颜色组合都能实现。图 4.24 表示了两种不可能发生的情况。在第一个情况中需要将夸克变成反夸克，许多原因都造成该过程不可能发生，最明显的是该过程中电荷不守恒，因为胶子不会将电荷带走。在第二种情况中整体颜色不守恒——输入的是 r 和 b，但输出的是 g 和 b，因此过程也不可能发生。

在 QED 理论中，电荷同性相斥、异性相吸，QCD 理论中的颜色荷也有类似的性质。具有同样颜色的两个夸克相互排斥，但

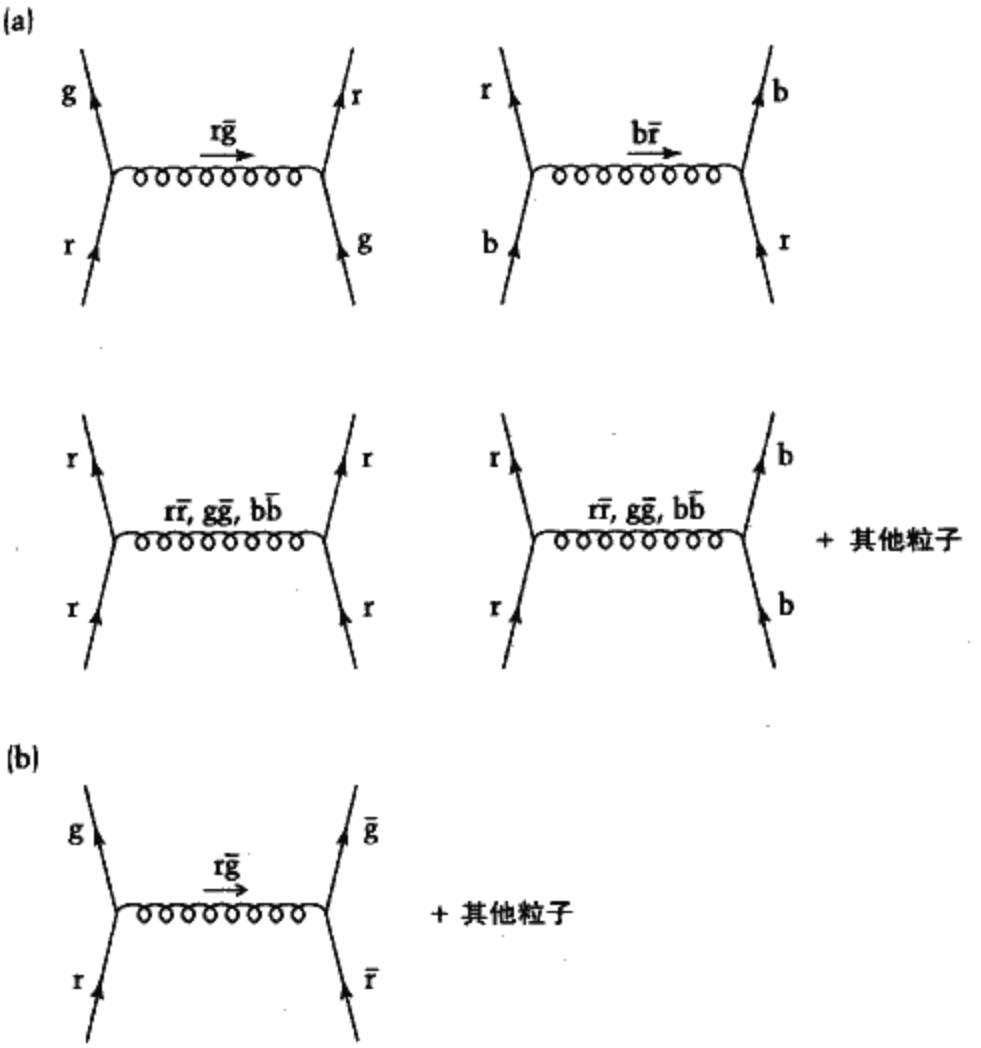


图 4.23 夸克-夸克相互作用［图 (a)］和夸克-反夸克相互作用［图 (b)］中的几种胶子颜色可能性

具有不同颜色的夸克相互吸引，除非发生交换颜色的过程——在这种情况下两夸克是相互排斥的。带有某颜色和与之对应的反颜色的夸克与反夸克会相互吸引，但如果反夸克所带的反颜色并不与夸克所带的颜色对应，它们就会互相排斥，这些都在图 4.25 中作了说明。

任何夸克和胶子之间的基本相互作用强度都是相同的。这并

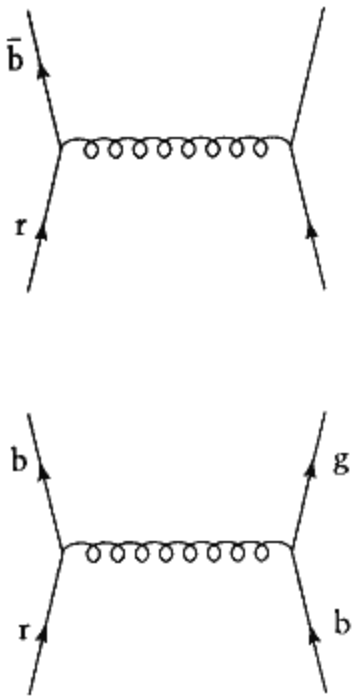


图 4.24 禁止发生的情况：第一种情况不满足夸克-反夸克守恒，因而不可能发生；第二种情况不满足整体颜色守恒

不是一种引入 QCD 理论的假设，而是一种被称作“普遍性”的理论特征——夸克同强子之间的基本耦合强度不会随夸克和胶子的颜色不同而发生改变。

为了说明这一点的重要性，我们有必要将 QCD 理论同 QED 理论再次加以比较。在 QED 理论中，带电粒子同光子发生相互作用的强度仅仅由其所带电荷决定，其电荷值可能是 1、2 或者电荷单位的其他整数倍。就像在第三章中所谈到的，QED 理论中并没有限制电荷的因素，也没有说明为什么它必须以某个基本单位的整数倍出现。而在 QCD 理论中，任何胶子以及任何夸克之间，相应的基本耦合强度都具有同一个值（但没指明是哪个值）。这样，一个红色夸克同胶子耦合的强度，与一个绿色夸克或蓝色夸克同胶子耦

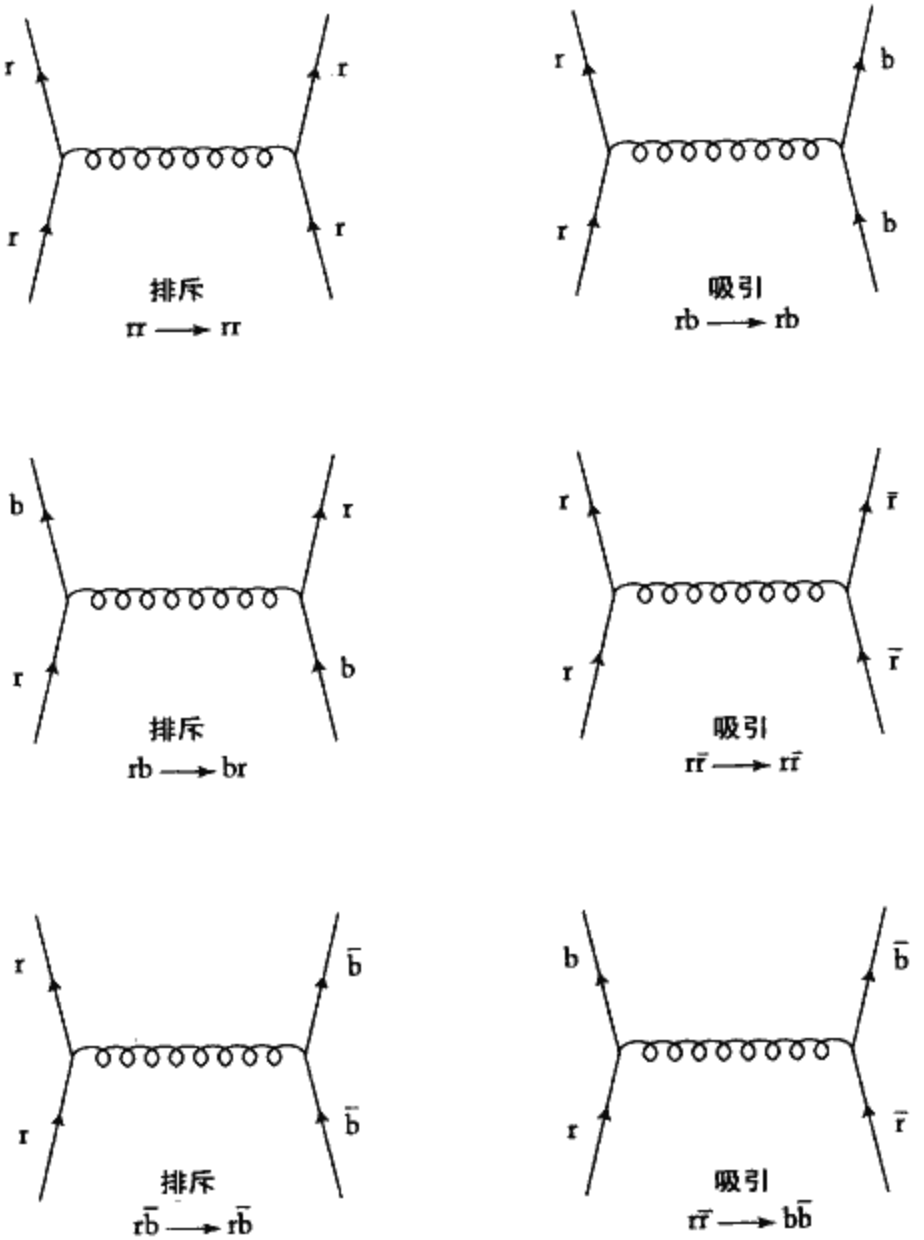


图 4.25　夸克-夸克和夸克-反夸克之间的吸引和排斥

合的强度完全一样。此外，每个夸克所携带颜色荷的值只能是一个颜色单位：如果一个夸克所携带的颜色值是两个单位而不是一个，它同胶子的耦合强度就会变为原来的两倍，但这是不允许的。如果实际情况不是这样，规范不变性就会遭到破坏，QCD 理论也会土

崩瓦解。

QED 理论同 QCD 理论的这种差别也可以归咎于 QCD 理论的非阿贝尔本质。从基本操作的角度来看，非阿贝尔群中，两个群的操作以某一顺序进行，和以相反顺序进行是有差别的，该差别可以用一个单操作乘上某个常数来表示。最重要之处在于，它确定了单操作和对操作①之间的关系，因此会存在许多限制。普适性正是这些限制的直接产物。非阿贝尔性看起来可能是一个很简单的属性，但实际上，它却将不同次数的变换联系在一起，在以上例子中出现的次数是一次和两次，这就意味着杨-米尔斯理论是那种被数学家们称为“非线性”的理论。孩子们是非线性的：拥有两个孩子的家庭同拥有一个孩子的相比，并不是简单地承受两倍的痛苦，或者拥有两倍的快乐，因为两个孩子之间会有互相作用，最终产生一种复杂得多的情形。像孩子的情况一样，非线性数学问题实际上是无法处理的，QCD 理论肯定也不例外。与之相反，物理上大多数较容易处理的问题，都是线性问题——它们只包含同样次数的变换或变量。图 4.26 阐释了一个小例子，说明一个线性系统很容易变成一个非线性系统。

QCD 理论中的非线性是从 SU（3）群的非阿贝尔本质而来的，这使得它很难在数学上进行处理，但这也正是 QCD 理论的诱人之处。在现实中，它表现为胶子本身也携带颜色，因此胶子之间能够直接相互作用。QCD 理论的主方程确实也可以通过费曼的路径积分表示，然后再通过微扰论生成 QCD 理论的费曼图，但最终出现的却是两张完全陌生的费曼图，在 QED 理论中不存在与之类似的费曼图。图 4.27 所示的两张图分别是三胶子图和四胶子图，它们清楚地展示了胶子之间的相互作用。作为普遍性

① 译者注：这里的对操作指上文提到的以某一顺序进行的一对操作。例如，对操作 A、B 依赖于它们的顺序，AB 与 BA 并不相同，其差别 $AB-BA$ 可以用另一个单操作 C 表示，即 $AB-BA=kC$，k 是一个常数。

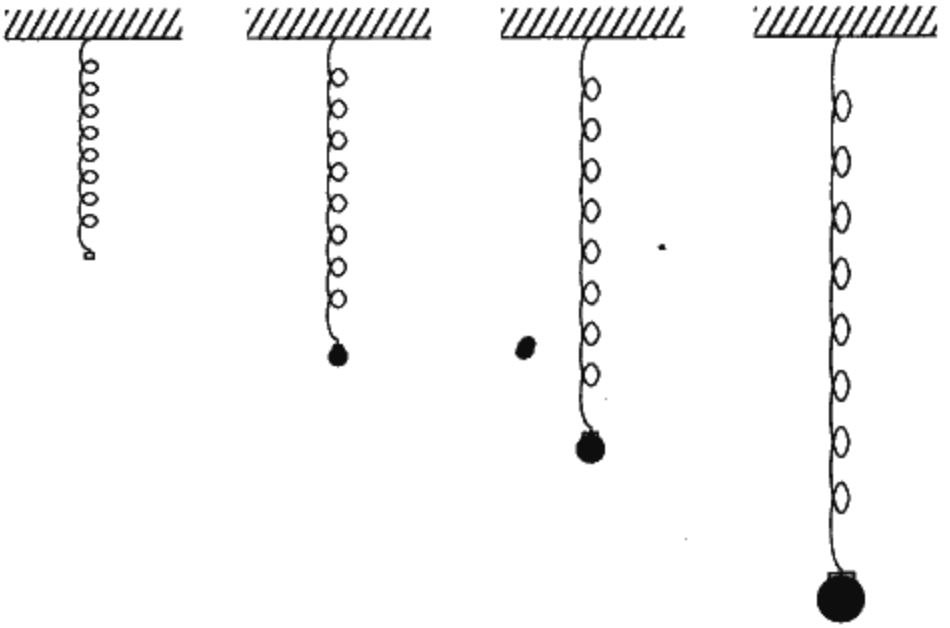


图 4.26　在弹簧的一端拴上一个小质量物体，弹簧就会从原来的位置伸长一点，即从最左边那幅图变为第二幅图。如果将物体质量增加一倍，弹簧伸长量也会增加一倍，见第三幅图。这是一种线性关系——质量加倍，伸长量也加倍。但是如果物体的质量太大，弹簧被拉坏，也就是最右边那幅图，简单的线性关系就会遭到破坏，系统就会变成非线性系统

的结果，三胶子顶点处的耦合强度同与夸克和胶子的耦合强度相同，四胶子顶点处的耦合强度是基本夸克-胶子耦合值的平方。

色、荷和流

两个电子之间的作用力源自于它们所带的电荷。很容易确定，电荷就是诺特定理所给出的那种著名的守恒荷，并且它很明显就存在于物质粒子上。在 QCD 理论中，人们会想到色荷与电荷类似，也是颜色力的起源，但是我们还需要从新的角度来考虑问题——携带力的粒子本身也带有颜色，量子标签数和守恒荷之间的搭配并不一致。怎样把这些同 QED 理论中费米子携带守恒

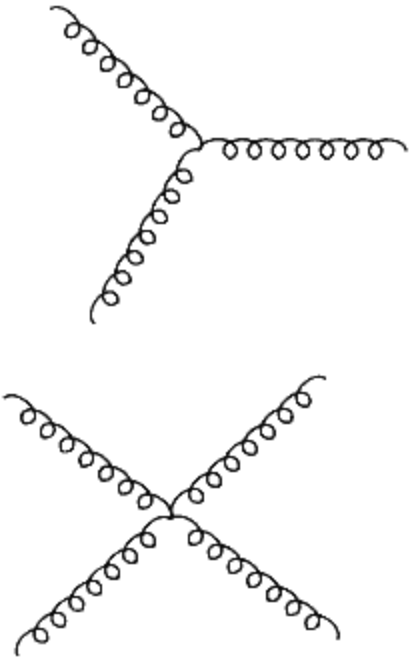


图4.27 三胶子和四胶子顶点。这些都是胶子的自相互作用费曼图

荷的简单情形联系起来呢？

QCD理论的基础是三种颜色，而不像QED理论中那样只有一种电荷。QCD理论中存在三种反颜色，这同存在正负电荷相类似。但除了替换掉QED理论中唯一的守恒荷之外，QCD理论还提供了更丰富的内容：群的数学表示要求总共要存在8种与SU（3)群相对应的守恒荷。每种胶子都携带着一种守恒荷。

现在已经存在8种荷了，但量子力学规则不允许一个粒子同时具有8种荷，就像禁止同时标明一个量子系统的位置和动量一样。实际上，群论法则表明，在QCD理论中粒子能够同时携带的最大标签数是2个。因此QCD理论中的8种守恒荷，不可能和两个独立的守恒量子数一对一地对应起来，因为没有足够的量子数供它们分配。这与QED理论不同，QED理论中的情况很简单——一个量子数对应着一种守恒荷。

这没有什么大惊小怪的。毕竟颜色群SU（3)$_c$群只是一种

SU（3）群的特例，我们已经看到过 SU（3）群怎样分配两个标签了。回忆一下 $SU(3)_f$ 群的夸克模型，它就建立在两个标签——超荷和同位旋第三分量的基础上，它们能够在给定的家族中很好地标定粒子。例如，所有重子八重态中的重子和介子八重态中的介子，都能够在一张超荷-同位旋第三分量的图上表示出来，见图 4.14。

仿照夸克模型，$SU(3)_c$ 情形下所用到的两个标签，可以叫做色超荷和色同位旋第三分量。因为可观测的粒子都不带颜色，因此不存在与之相应的日常经验，所以给这种永远都不可能看到的量子标签起名字，根本没什么头绪。但很明显它有类比的对象，可以严格效仿用普通超荷和同位旋第三分量标记介子八重态的方式，用色超荷和色同位旋第三分量标记 8 种胶子。介子存在于从味道 $3\otimes3^*$ 产生的八重态中，而胶子存在于从颜色 $3\otimes3^*$ 产生的八重态中，这两个群都可以用各自的量子数来表示，见图 4.28。出现在介子八重态中间的两种粒子是 η 粒子和 π^0 粒子，其超荷和同位旋第三分量的值都是 0，它们分别对应于 $\frac{1}{\sqrt{6}}(r\bar{r}+g\bar{g}-2b\bar{b})$ 和 $\frac{1}{\sqrt{2}}(r\bar{r}-g\bar{g})$ 这两种胶子颜色组合方式。其他六种胶子颜色组合方式形成了图 4.28 中胶子八重态的边界。

在经典的电磁学中，点电荷会产生电场。运动电荷形成了电流，电流也是守恒的。电流既能产生电场也能产生磁场，这在量子理论中就变成了电流能产生光子的思想。将同样的想法运用到 QCD 理论中，可能会得到比寻找 8 种守恒荷更丰硕的成果，即寻找 8 种相应的守恒流。

在经典电磁学中，当电荷的运动方向发生改变，就会发出电磁辐射，描述辐射传递的定理就会起作用。从量子角度上讲，发射光子与带电粒子运动状态的改变有关，图 4.29 用图表的形式

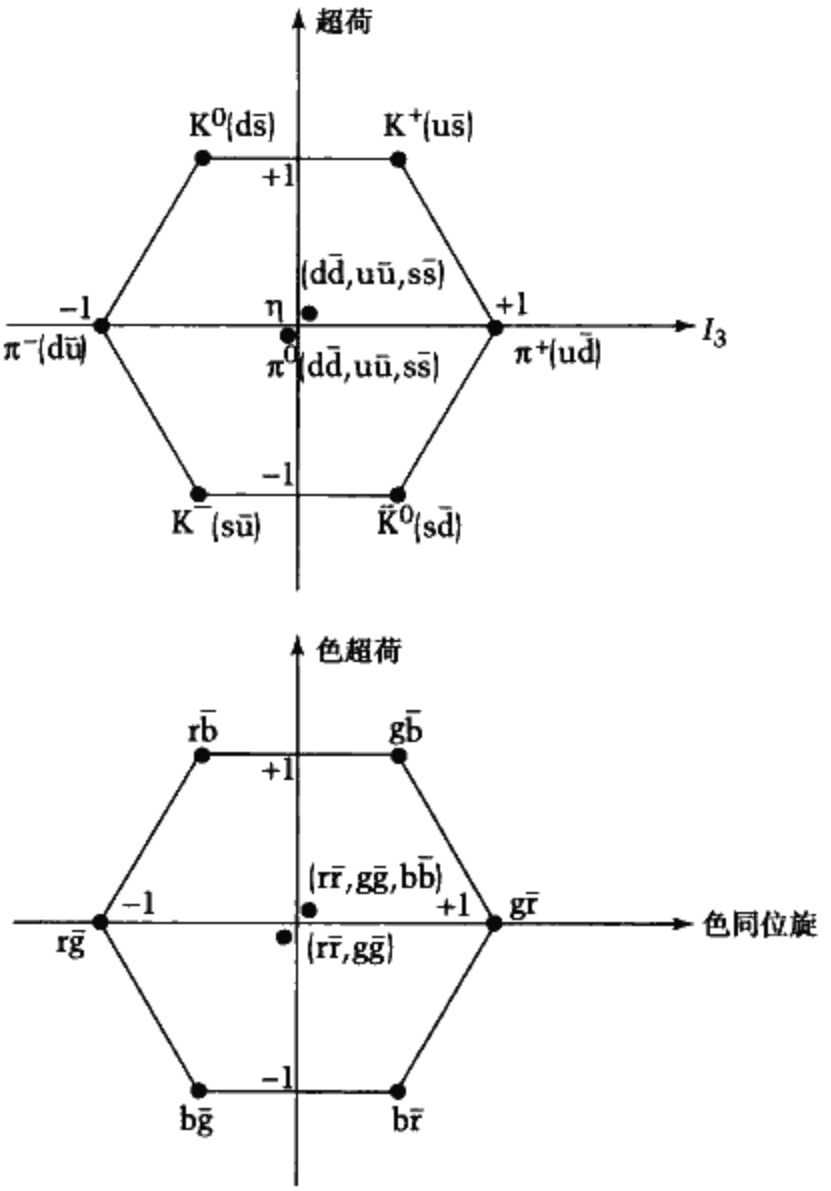


图 4.28 胶子色八重态（下面的图）可以用类似于介子味八重态（上面的图）的方式进行标记

表现了这个过程。由于光子会带走四维动量，电子的四维动量也会发生改变，图中的四维动量用 k 和 h 表示。由于光子不可能带走电荷，因此电子的电荷不发生改变。图中虚线框表示守恒区域，进入该区域的四维动量的值等于离开区域的四维动量值，并

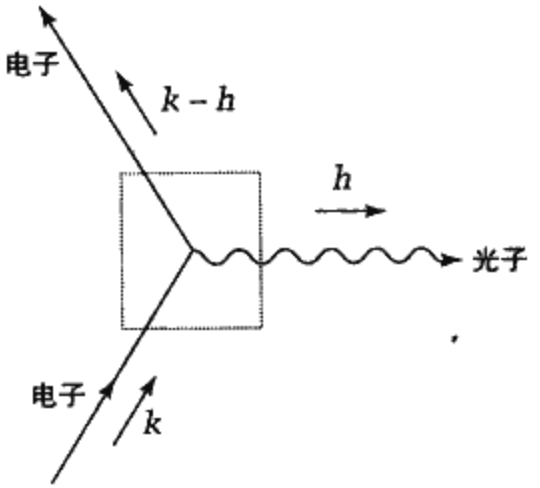


图 4.29 一个动量为 k 的电子放出一个动量为 h 的光子，这样，它就变成了一个动量为 $k-h$ 的电子

且进入该区域的电流也等于离开区域的电流。

带颜色的夸克会以类似的方式放出胶子，图 4.30 的图（a）形象地将该过程表示了出来。图中的点划线表示颜色流，表示颜色流的线不能在这种图中起始或终止。就像电子在放出一个光子之后要保持电荷不变一样，夸克的颜色——如果它看起来是红色的话，其实它是绿色或蓝色的也一样——也可以不发生改变。但夸克的颜色也能够改变，如图 4.30 的图（b）所示的情况。图 4.30 的图（c）是一个更复杂的例子，图中包含一个三胶子顶点。在这三种情况中，向矩形框所围区域内流入的颜色流都等于流出的颜色流。存在 8 种守恒颜色流，可以用 8 种有效的颜色结合方案（即 8 种胶子）对其进行有效标记，它反过来又对应着与胶子发生相互作用时，夸克颜色演化的 8 种方式。

在 QED 理论中，只存在一种能产生光子的守恒电流；而在 QCD 理论中，存在 8 种产生胶子的守恒色流。在所有的规范理论中，与对称性相关的流都是携带力的场的来源。

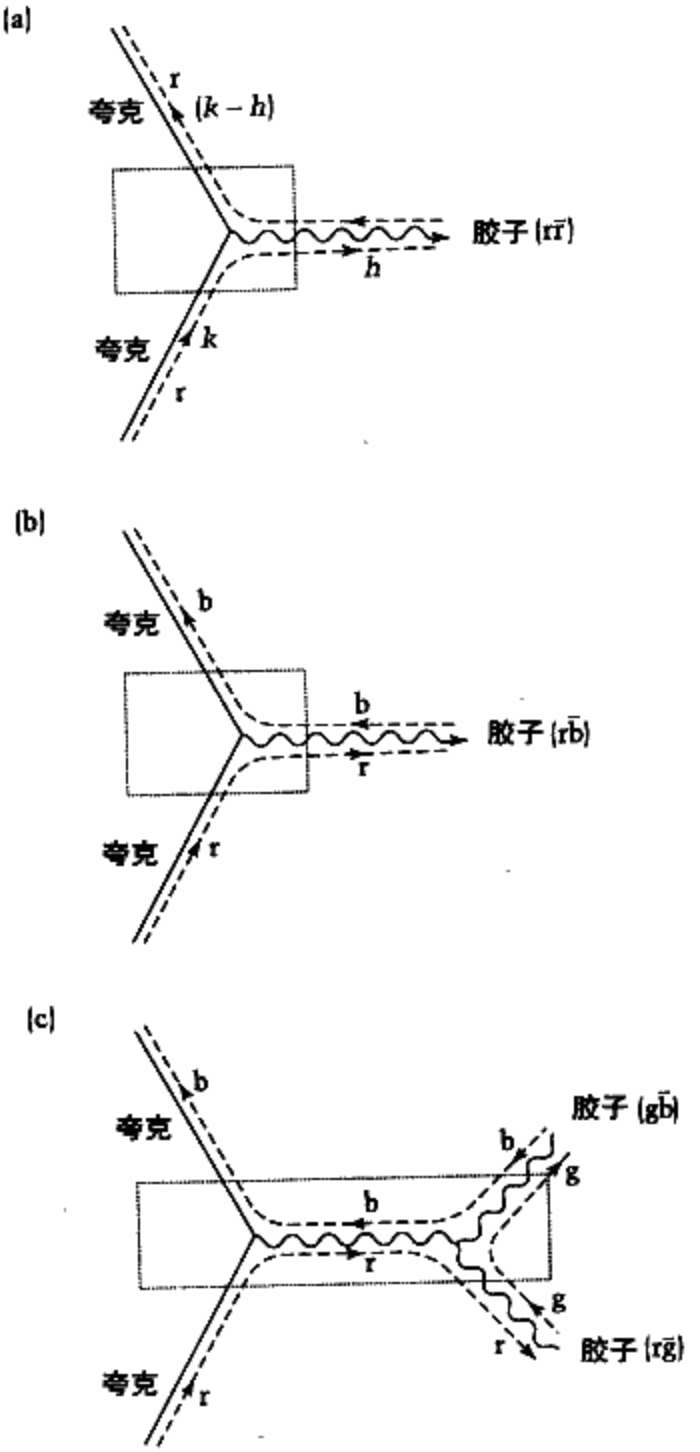


图 4.30 夸克-胶子和胶子-胶子相互作用中的颜色流

存在颜色的证据

如果没有颜色属性，就不会有 QCD 理论，这是一个简单明了的事实。因为根据 QCD 理论，颜色是粒子的一种性质，粒子只有通过颜色才能发生相互作用。这就表示，实验者们要拿出一些十分确凿的证据来支持颜色的存在。不仅要这样，他们还必须证明正好存在 3 种颜色，不能多也不能少。

实验物理学家们已经接受了挑战。他们已经提供了颜色存在的证据，或者至少发现了一些人们预测颜色所应该具有的行为。他们也发现颜色的数量恰好是 3 种的证据。他们甚至还更进一步，追踪到了对称群本身的精确形式。

一些实验进行得更为直接。可以论证，所有方法中最直接的就是让电子和正电子迎头相撞发生湮灭，然后检查产生的 μ 子数目及相关的强子数。

先等一下。这是电子-正电子碰撞的过程，从哪儿来的强子呢？即使在今天，电子和正电子仍然被认为是最基本的粒子。因此它们并不是复合粒子，并且它们也不参与强相互作用。实验中出现强子的现象是对狭义相对论和量子场论的检验。从一种意义上——狭义相对论的意义上来讲，碰撞粒子总质量加上碰撞粒子的动能在碰撞时产生了大量能量，这些能量在粒子之间进行分配，从而产生了许多粒子碎片。从另一种意义上——量子场论的意义上来讲，以前所不存在的粒子是能够产生的，其中一些会发生强相互作用，只要不违背基本的守恒定律就可以。

图 4.31 用简单的费曼图表示了两个过程，一组高能电子-正电子对产生了一个能量非常高的虚光子，光子自己又产生了一组带电的费米子-反费米子对，该过程在效果上同最初的碰撞过程相反。光子确实也可能又产生一组电子-正电子对，但这种过程在实验中存在的大量电子-正电子对之间是很难观测到的，并且

实验中许多发生相互作用的电子-正电子对实际上不发生湮灭。而另一方面，μ 子-反 μ 子对必须通过湮灭过程产生，因此它也是电子-正电子对发生湮灭的标志。通过同样的方式，夸克-反夸克对也可以由这种湮灭产生。但与 μ 子和反 μ 子不同，这些夸克与反夸克不能够在实验者的探测器上被直接观测到。事实上，这些夸克和反夸克会在探测器上产生大量的粒子闪光。

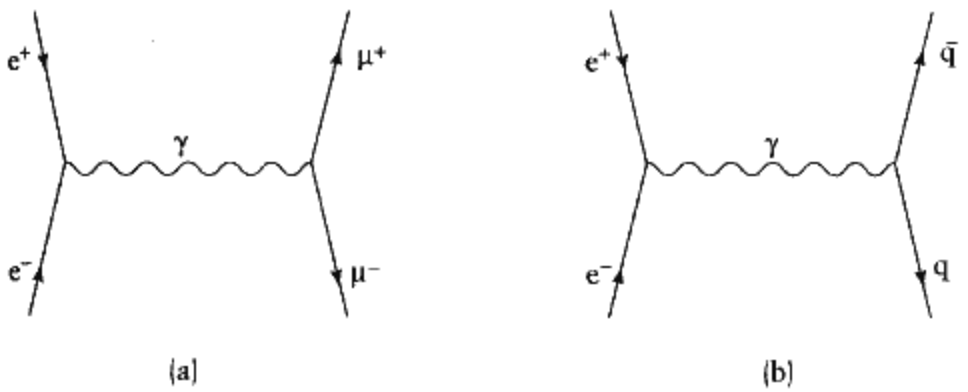


图 4.31 一个电子和正电子发生湮灭，可以生成一个 μ 子-反 μ 子对［图 (a)］，或者一个夸克-反夸克对［图 (b)］。每个夸克或反夸克又能够产生一次强子喷注

实际上，如果存在 3 种夸克味道，就可能出现 3 种不同的夸克-反夸克组合：u 加上 $\bar{u}$、d 加上 $\bar{d}$ 以及 s 加上 $\bar{s}$。这 3 种味道的夸克所带的电荷（u、d、s 三种夸克所带的电荷分别为 $+2/3$、$-1/3$和$-1/3$），控制着它们同传递相互作用的虚光子之间的电磁耦合强度，而这些虚光子又能喷射出夸克-反夸克对。由于 μ 子和反 μ 子所带电荷分别是 -1 和 $+1$，因此与产生 μ 子-反 μ 子对相比，产生夸克-反夸克对的概率分别是 $(2/3)^2$、$(-1/3)^2$ 和 $(-1/3)^2$，也就是 4/9、1/9 和 1/9。这反过来又表示，产生的夸克-反夸克对同 μ 子-反 μ 子对的比率是 $4/9+1/9+1/9=6/9$。

实际情况不是这样。这是因为在计算过程中漏掉了一些东西，那就是颜色。对于 3 种味道的夸克，每一种味道又具有 3 种

可能的颜色，产生的每种夸克-反夸克组合的可能性都是原来所预料的 3 倍，因此总几率必须用 3 乘以上面的数，得到 18/9，也就是 2。

如果说这就是实验者们所观测到的现象，一切都会显得干净利落，但这也不是实际的情况。在能量不到 3 吉电子伏的电子-正电子对撞实验中，测得的比率的确大约是 2，尽管这已经很不容易，实在叫人难以置信了。但如果用比率对能量作图，就会得到一个逐渐上升、带有一些起伏的曲线，最终在大约 10 吉电子伏的地方变成了一条平坦的直线。从这里一直到大约 45 吉电子伏的地方，比率很明显一直是常数，之后比率值大幅上升。在这段很长的平坦部分，比率是 11/3，而不是 2，如图 4.32 所示。其中比较好的消息就是，比率图中确实存在着一个“平台区”，但坏消息就是其对应的值不正确。出现这种明显偏差的最好解释就是：还存在为几率作贡献的另外两种夸克，其中一种电荷为＋2/3，叫做粲夸克（charmed quark），于 1974 年被发现；另外一种所带电荷为－1/3，叫做底夸克（bottom quark），于 1977 年第一次被观测到。用这 5 种夸克，并且应用以上算术思想，也就是对电荷平方求和，然后将结果乘以颜色因子 3，就得到了 11/3。这同能量在 10 吉电子伏以上“平台区”阶段的实验值一致——实验支持了 3 种颜色和 5 种夸克的存在。

加速器的插曲

在图 4.32 所示的比率图背后，隐藏着大量的实验努力和硬件支持。其中，许多数据都是由众多实验小组利用 20 世纪 70 年代和 80 年代那一批电子-正电子对撞机得到的。其中第一台机器，就是位于 SLAC 的斯坦福正负电子非对称环（Stanford Positron Electron Asymmetric Ring），或简称为 SPEAR。SPEAR 的周长为 234 米，最终能使具有 4 吉电子伏能量的电子同 4 吉电子

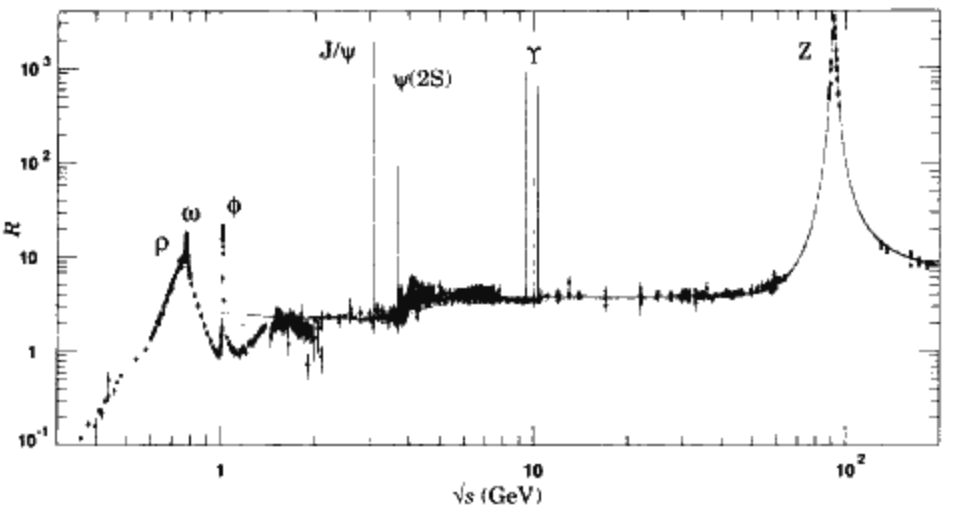


图 4.32 比率 R 是电子与正电子碰撞产生强子的散射截面与电子与正电子对撞产生 μ 子-反 μ 子对的散射截面之比。J/ψ 粒子、Υ 粒子以及其他态都能在图中显示出来。在 J/ψ 粒子峰出现之前，比率值是 2，而不是 6/9。因子 3 的差别是存在 3 种颜色的证据。巴丁、弗里奇和盖尔曼在 1972 年的一次会议讲座中首次指出这一点，他们还在这次会议上讨论了中性 π 子衰变过程的颜色问题。在 Υ 共振态峰以上，R 值跳到了 11/3；如果不存在颜色属性，这个值将会是 11/9。图片来源：粒子数据小组

伏的正电子发生对撞，碰撞总能量达到 8 吉电子伏。许多在有关 QCD 理论的故事中出现的特征都是由 SPEAR 发现的。

SPEAR 最早期的任务之一，是测量强子产量与 μ 子产量的关键比率，但它达不到很高的能量。为了达到更高的能量，接力棒传给了 SLAC 对撞机——它被称作正负电子对撞机（Positron Electron Project collider），或简称为 PEP。这台机器在 20 世纪 80 年代开始运行，其周长为 2.2 千米，电子同正电子对撞的总能量达到 30 吉电子伏。PEP 在运行 10 年之后停止了工作，让位给耗资 1.77 亿美元建造的升级版机器，这台新机器叫做 PEP Ⅱ，或者可以更形象地称之为非对称 B 粒子工厂（Asymmetric B-Fac-

tory)，它在 1999 年投入使用。这台机器可以用来研究一种很重的介子，叫做 B 介子。

在 SPEAR 和 PEP 之间（无论是在时间顺序还是能量范围上）还存在着 CESR——位于纽约州伊萨卡的康奈尔电子-正电子存储环（Cornell Electron-positron Storage Ring）。CESR 周长约为 768 米，它能将亮度很高的 8 吉电子伏粒子束注入一个巨大的粒子探测器，使之发生对撞。这个探测器有一个令人愉快的名字叫做 CLEO，是为了与 CESR 配对而起的，它们分别借用了恺撒（Caesar，罗马恺撒大帝）和克利奥帕特拉（Cleopatra，埃及艳后）的名字。

但美国绝不可能垄断电子-正电子对撞机市场。在欧洲运行着 DORIS 和 PETRA，两台机器都位于德国的加速器中心 DESY。DORIS 即正负电子双存储环，也叫做 DOppel-Ring-Speicher（双存储环，Double Storage Ring），是这两台机器中较先启用的，也是较小的一台，它在 1974 年投入使用，DORIS Ⅱ 的能量可以达到 11 吉电子伏左右。而在 1978 年开始运行的 PETRA——正负电子环状队列加速器（Positron Electron Tandem Ring Accelerator）更加庞大，它能够将 23.4 吉电子伏的电子与相同能量的正电子发生对撞，其总能量达到了 46.8 吉电子伏。PETRA 还会在 QCD 的故事中再次出现，这当然不单单因为它是一座仍然在进行粒子物理研究的加速器（现在它已经经过改装了）。

粒子加速器不会被轻易遗弃，一些旧机器通常被当作注入器或辐射源来使用。“注入器”就是一种小型加速器，它能为更大的加速器提供预加速的粒子束。PETRA 现在就扮演着这种角色，它为更大的 HERA 电子－质子对撞机提供预加速，HERA 是 1992 年开始在 DESY 投入运行的。同时，DORIS 经过多次升级后，变成了一座强大而又容易控制的 X 射线源，它被用于医药、材料、化学等多个领域的研究。欧洲分子生物实验室在 DESY 设

有一个站点，专门开发这方面的用途，这或许能成为粒子物理学研究的一种惊人的副产品。

综上所述，实验物理学家们利用这一批机器，在电子-正电子对撞过程中得到了许多实验数据，这些数据给出了十分可靠的μ子-反μ子对产量同强子产量的比率。该比率告诉物理学家们存在5种夸克，并且夸克的颜色有3种。但是以CERN的LEP为先锋的一批新型电子-正电子对撞机提供了更大的能量，如此高的能量为改变这一经典结果提供了可能。

LEP——世界上最大的加速器

没有人会大批量生产巨大的粒子加速器。每一台加速器都是定制并手工建造的，并且每一台都获得了成功，这成为了现代工程史上的奇迹。加速器之所以如此引人注目，是因为它能够将最大的尺度和最高的精度结合在一起。

我们可以以LEP——位于CERN的大型电子-正电子对撞机（图4.33）为例了解一下加速器。LEP位于法国-瑞士边境侏罗山（Jura）附近，深埋在地下一条周长为26.67千米的圆形管道中，直到2000年底LEP关闭的时候，它还是世界上最大的粒子加速器。LEP于1989年完工，造价达10亿美元，在20世纪80年代中一半的时间内，LEP都是整个欧洲最大的民用工程。这是真正的大科学项目。

LEP隧道是一项极其精确的地下工程。隧道所在的平面有1.4°的倾角，以确保尽可能多的隧道穿过坚硬的石层，而不是松散的泥土。工程师们通过激光干涉技术保证整个27千米长的圆周周长误差小于1厘米。这是一个很大的地洞。在地下50米到100米之间要挖开一条直径约4米的隧道，还有一些实验洞穴和竖井，必须挖出1400万立方米的沙土和石头。在地上还要建造70多所建筑，其中一些安放着大约160台进行控制的计算机和微

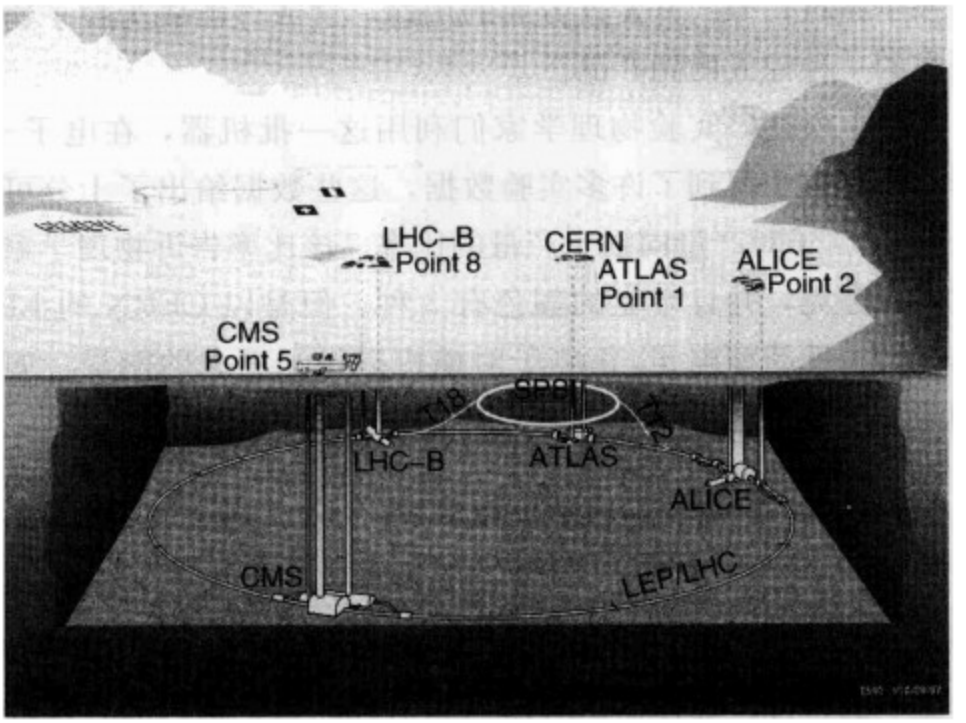


图 4.33　地表之下的怪物。位于法国-瑞士边境地下的 LEP 隧道，现在被大型强子对撞机（LHC）所占据。圆环上的凹槽是进行实验的区域，LEP 的实验物理学家们占用了其中的 4 个：L3（地点 2）；ALEPH（地点 4）；OPAL（地点 6）和 DELPHI（地点 8）。更新的 LHC 实验名称，已经在图上标注了出来，它们分别是：ATLAS（地点 1）、ALICE（地点 2）、CMS（地点 5）和 LHC－B（地点 8）。图片来源：CERN

处理器。

在隧道中，有 4000 多块磁铁用来控制粒子束，磁铁中间是总共 22 千米长的真空室。一些功率相对较低的磁铁被用来控制粒子束在一个大致为圆形的轨道中运行。其他磁铁通过使粒子束交替汇聚和发散的方式，形成交变梯度聚焦效应。当操作者用磁铁控制粒子束时，有 750 多个直流电源为设备供电，电源的功率从 1 千瓦到 7 千瓦不等，其精度要达到百万分之二十，而且还有 500 个传感器在监视着粒子束的位置。

加速粒子所需的交变电压通过射频电磁波来实现。因为电子和

正电子带有相反的电荷，同样的电压将会使它们在加速器环中朝着相反的方向加速。用来加速的射频场被施加在圆环某些特定的点上，如果不计磁铁所施加的力，电子在这些点之间是自由飞行的。在原始的LEP设计中，射频能量通过128个谐振腔传给粒子束。这些2米长的铜室包围着粒子束，它们通过16个1兆瓦的速调管（射频电磁波发生器）将能量传给粒子束。

LEP是一台电子-正电子同步加速器，它的正常工作要依赖于磁场及加速电压频率改变时的相位稳定性。这导致它不可能加速冷态的粒子，实验者不得不先将粒子预加速，然后再注入LEP中，现代加速器通常都是这样运行。

在LEP中，最初的电子束来自于加热的灯丝，就像在电视机显像管中发生的过程一样。正电子束由能量为200兆电子伏的电子束和适当的靶碰撞产生。随后，两种电子束都要通过600兆电子伏的直线加速器——直线加速器，也叫作线性加速器（linac），再继续进入600兆电子伏的存储器进行存储。粒子从这里转移到能量为3.5吉电子伏的CERN质子同步加速器（PS），紧接着它们又会进入另一个加速器——CERN超级质子同步加速器（SPS）。在电子和正电子束注入LEP之前，自身已带有20吉电子伏的能量。PS和SPS都经过改装，以便使它们能同时加速电子和正电子。在SPS中，电子-正电子的加速周期被穿插在质子加速周期之间，因此SPS在担当LEP注入器的同时，仍能够实现固定靶质子加速器的功能。要想将粒子加速到很高的能量，并不是每天早上打开加速器，然后坐几分钟等着粒子的速度快起来这么简单，见图4.34。

需要注入预加速粒子仅仅是LEP以及其他加速器（例如Tevatron）运转的必要步骤之一。另一个问题就是粒子并不以连续粒子流的方式到达，而是一团一团涌过来。在LEP，每一团粒子只有1.8厘米长，并且只有几团电子朝着一个方向绕转，而同样

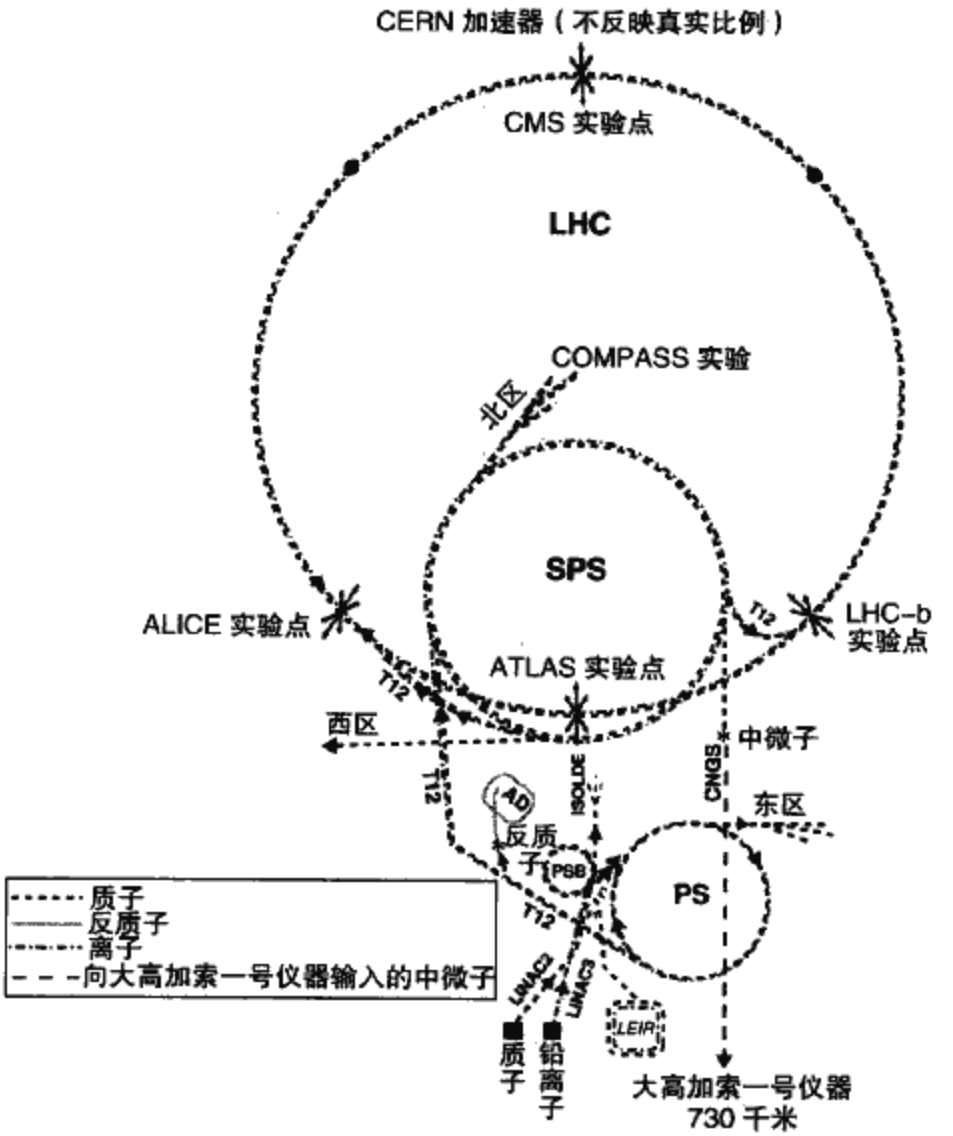


图 4.34　CERN 加速器的整体示意图。LHC——大型强子对撞机；SPS——超级质子同步加速器；AD——反质子减速器；ISOLDE——在线同位素质量分离器；PSB——质子同步推进器；PS——质子同步加速器；LINAC——线性加速器；LEIR——低能离子环；CNGS——CERN 中微子到大高加索计划。[1] 图片来源：CERN

[1]译者注：将中微子发送到意大利大高加索（Gran Sasso）的一个研究项目，用来研究中微子振荡。

数量的正电子团朝另一个方向绕转。可以证明，LEP 最令人吃惊的数据，就是电子团和正电子团以接近光速的速度——大约 30 万千米/秒的速度相互逼近，然后正好于安装在环上的四个探测器中心处相遇。在探测器中心处，每过几微秒就会发生电子团和正电子团对撞的过程。

粒子束越密集，实验者就能在探测器中看到越多的碰撞，因为粒子发生碰撞的机会更多。一台加速器的有效密集度称作它的亮度（luminosity）。亮度是一个加速器工程师最期望提高的量，这是改进加速器物理输出的最重要方法。在 LEP，碰撞点周围的特殊磁铁将粒子束挤压到一个大约 8 微米高、200 微米宽的范围内，这一范围越小越好，因为这样会有更多的粒子挤进更小的区域。这说明在碰撞点处，每团粒子截面的面积和针尖差不多大，但两团粒子还是能撞到一起。另一种提高亮度的方法，就是增加每团粒子的数量。例如，如果每团粒子中有 41.6×10^{10} 个电子或正电子，记录到的 LEP 亮度就是 11×10^{30} 个粒子每平方厘米每秒，这个数字还在逐年增长。实验物理学家们经常一连做几个星期甚至几个月的实验，以收集他们所需的数据，这样就产生了一个累积的亮度，叫做总亮度，它是亮度对时间的积分。举个例子，进行一次 LEP 实验的典型总亮度，可能是 35 个皮靶恩倒数，1 皮靶恩就是 10^{-12} 靶恩，也就是 10^{-40} 平方米。这个值等于两束粒子以上面所述的亮度值不间断地碰撞 37 天得到的结果。

人们通常认为，电子只能和正电子相互碰撞，而不能同其他物体碰撞。为了确保这一点，粒子束必须尽可能在真空中运动，只有这样，粒子束在一圈一圈绕着环状轨道的漫长旅途中，才没有机会同空气或者其他气体分子碰撞。在如此大的范围内维持真空状态是一件很困难的事情。真空泵从粒子飞行的真空管道中吸走了大部分空气。在 LEP 中有一条 3 厘米宽、22 千米长的锆铝合金条，它吸收了残留的个别杂质原子，这同用粘蝇纸粘苍蝇的

过程差不多。人们要定期“清洗”这条金属带，实验者们把它加热到400℃，这样杂质原子就会在金属带中弥散开，埋在里面永不翻身。

除掉了垃圾原子之后，电子和正电子就能够顺利对撞了。1989年7月14日，LEP刚一开始工作，就实现了该对撞过程，比预计时间提前了1天。在4个月的时间内，实验者们已经在LEP上捕获了10万多个Z粒子。这个数字确立了LEP作为“Z粒子工厂”的地位，它能够产生大量的Z粒子，并且能很细致地研究Z粒子的性质，因此这成为LEP的主要研究计划之一。

最初议程中的另一项条款，是最终要将LEP升级到更高的能量，从而能够产生W粒子对。Z粒子之所以可以单独生成是因为它是电中性的，从电子和正电子碰撞过程产生单独的Z粒子不会受到电荷守恒定律的制约。每束粒子具有50吉电子伏能量，或者说在碰撞总能量为100吉电子伏的条件下，原始的LEP已经能很好地产生质量为91.2吉电子伏的Z粒子。而带电的W粒子必须以W^+和W^-粒子对的形式产生，要想看到质量为160.6吉电子伏的W粒子对，LEP必须将能量提升到能看到W粒子对产生的范围。

这是将LEP升级为更高能量机器（称作LEP2）的主要目的。为了达到这一能量，工程师们将铜制的旧加速腔换成了新的超导材料。到1996年7月，144个超导腔已经安装到机器上，LEP2的总碰撞能量达到了161吉电子伏。在这种更高的新能量下，启动机器2天之后，LEP的实验者们见证了第一组W粒子对的产生。

能量的增加以及超导腔的应用带来了新的挑战。电子和正电子现在每秒能绕27千米长的加速器旋转大约11 200次，每转一圈都要损失2%的能量，这些能量要及时补充，以免粒子的速度慢下来。之所以存在能量损失，是因为任何加速运动的带电粒子

都会放出电磁辐射。在粒子加速器的能量下，这种所谓的同步辐射以X射线的形式表现出来。因为会产生大量的同步辐射，所以加速器有时修建在地下，大地本身就成为一个天然的辐射屏蔽器，LEP就是这种情况。此外，粒子束所在的真空腔也需要进行特殊屏蔽，这样辐射就不会破坏电缆及管道中的其他仪器。即使是这样，在LEP运转的时候，还是不允许有人进入隧道，因为辐射的危害太大了。但黑暗之中总有一线光明，我们可以将加速器产生的同步辐射转化为有用的X射线源，像DORIS和DESY那样用于医疗或材料方面的研究，但总的说来，辐射还是不受欢迎的。实际上，LEP环形轨道巨大的周长就是用来减少同步辐射的，因为轨道弯得越厉害，辐射就越强。但是能量的负面影响最为严重：如果加速粒子的能量变为原来的2倍，则同步辐射的损失就会增加为原来的16倍。而在1998年完成的升级就是要将LEP的能量增加为原来的2倍，其总的碰撞能量达到了192吉电子伏。正是同步辐射所带来的能量损失，限制了像LEP这种电子-正电子加速器的总功率。

被LEP2用来提高能量的超导腔也带来了新的挑战。在传统的铜腔中，由于腔壁电阻的热效应，震荡的电磁场会损失许多能量。但在超导腔中，用超导体铌所做的腔壁几乎不存在任何电阻。这使得超导腔的效率更高，并且能在高得多的电场下运行，但是这同样要付出代价——超导体一定要冷却到液氦温度，大约是−269℃。为了达到该温度，CERN必须生产和储存6万升液氦，这就意味着要建立一座世界上最大的液氦工厂。

LEP项目的进一步规划也包括了LEP的死亡计划：LEP在2000年停止运行，让路给大型强子对撞机（Large Hadron Collider），或者叫LHC，它将会占用LEP的隧道。LHC计划于2005年竣工，它能够使质子发生对撞，而不是更常见的质子-反质子过程，它将成为地球上最庞大、能量最高的加速器。

LEP 的主要竞争对手在大西洋彼岸，它是于 20 世纪 90 年代落成的另一座 Z 粒子工厂——斯坦福直线对撞机（Stanford Linear Collider），或称 SLC。SLC 于 1989 年——和 LEP 在同一年——启用，但比 LEP 早 2 年停止运行。一方面，具有 100 吉电子伏能量的 SLC 是最初那台位于 SLAC 的 3.22 千米（2 英里）长的直线加速器的升级产品，另一方面它又是一台完全与众不同的机器。

SLC 的基本思想是利用一台直线加速器同时加速电子和正电子。先从直线加速器中取出一些电子，反应得到正电子，正电子又被送回加速器的起始端。在主加速器运行之前，电子和正电子都要经过称作阻尼环的小存储环，它的作用是降低粒子束的弥散度。而在加速器中，每团正电子后面都紧随着一团电子，两个粒子团的时间差只有 60 纳秒（1 纳秒$=10^{-9}$秒）。后面还紧随着另一团电子，这团电子会在改变方向后离开直线加速器，然后被用来产生正电子。当这两套被加速的粒子离开直线加速器的末端，具有 50 吉电子伏能量的电子被导入一个弧形的弯曲磁铁，而具有 50 吉电子伏能量的正电子被导入另一个相反的弧形，最后这两束粒子迎面相碰。从 1991 年开始，一台精心制作的探测器——SLD，或者称为 SLC 大型探测器（SLC Large Detector）——开始检测和记录这些碰撞。

SLC 还是新一代电子-正电子对撞机的实验台。如果仅仅按比例放大 LEP 来制造能量更高的机器，机器会大得不可思议，并且会通过辐射损失太多的能量。其解决方案可能就是在一条直线上加速电子和正电子，然后让它们拐弯以迎面相撞。就像 SLC 那样。SLAC 的科学家和工程师们已经开始积极研究 SLC 的下一代机器，它被暂时地称作下一代直线对撞机（Next Linear Collider），简称 NLC。以现在的情况看来，LEP2 很可能会作为人类建造的规模最大、功率最强的回旋正负电子对撞机而永载史册。

OPAL的颜色

与早期那一批电子-正电子对撞机相比，能量更高的LEP和SLC开拓了一整片潜在的全新领域。

在低能情况下，电子和正电子能够湮灭成光子，而该光子转而又能产生一组电子-正电子对，或者是μ子-反μ子对，甚至还能够产生夸克-反夸克对。后两种情况在PETRA、PEP以及其他一些对撞机中，都被用作寻找颜色存在的证据。

但在LEP和SLC所具有的能量下，电子和正电子湮灭还能另外得到一个电中性的Z粒子，Z粒子又可以衰变得到一组费米子-反费米子对，例如电子-正电子对或者夸克-反夸克对。因此更高的能量提供了一个全新的反应道，电子、正电子可以经由Z玻色子发生湮灭。不仅如此，再注意一下夸克-反夸克可能发生的过程：每个夸克或反夸克在离开碰撞点的时候都具有极高的能量，因而它们极有可能反过来又抛出高能胶子——这是一种韧致辐射的形式。

如果每个夸克和反夸克都能放出一个胶子，那么一次电子-正电子湮灭有时会产生4个碰撞产物——1个夸克、1个反夸克和2个胶子。但是还有其他3种能够产生这4个粒子的方式。例如，夸克或反夸克本身都有可能发出2个胶子，或者由夸克或反夸克释放出1个胶子，该胶子又生成一组新的夸克-反夸克对。这4种过程的发生概率大致相同，图4.35是它们的费曼图，这只是一整套相关反应费曼图的一部分。

但是实验者们利用SLD或LEP探测器，并没有在碰撞产物中看到夸克、反夸克和胶子。取而代之的是，这些碰撞产物自己生成了发生强相互作用的剩余粒子，实验者们观测到的正是这些粒子组成的喷注，或者叫喷射。图4.35是这些过程的标志性特征，即明显的4条强相互作用粒子喷注，称作强子喷注，它与电

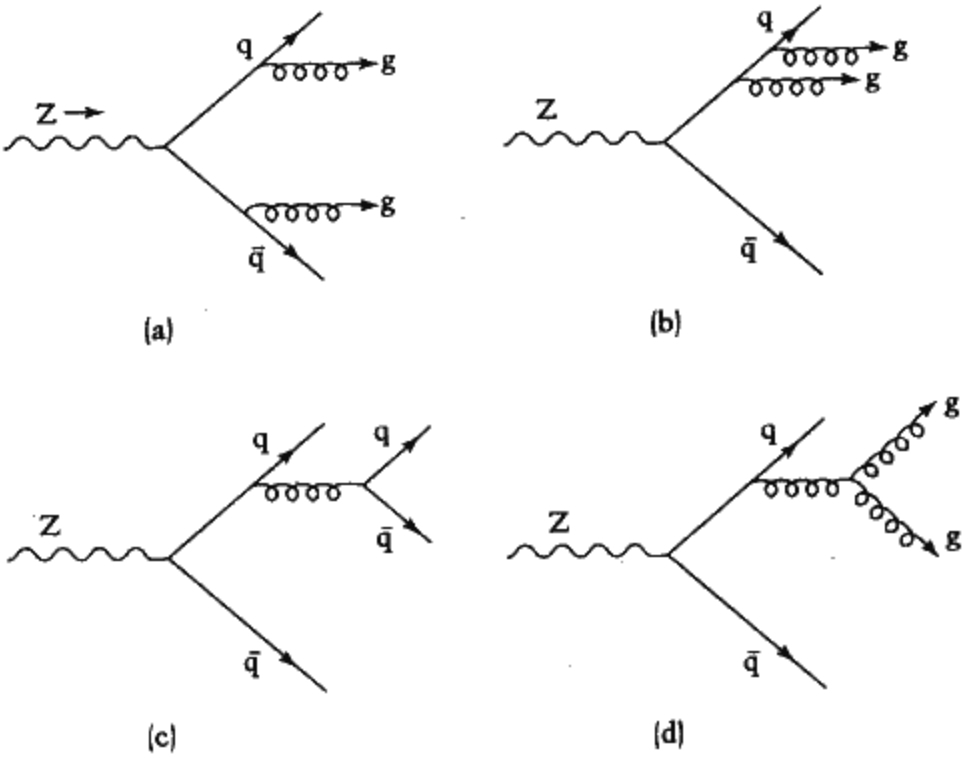


图 4.35 电子-正电子通过 Z 粒子发生湮灭，产生四喷注：(a) 和 (b) 都是双胶子韧致辐射，产生 $e^+e^- \to q\bar{q}gg$ 过程。图 (c) 是有粒子对产生的 $e^+e^- \to q\bar{q}q\bar{q}$ 过程。图 (d) 是包含三胶子顶点的 $e^+e^- \to q\bar{q}gg$ 过程

子-正电子散射截面中的 Z 粒子峰值一致。

这种经由 Z 玻色子产生的四喷注现象，为检验颜色群提供了一种重要的方法。当理论物理学家们计算散射截面的时候，他们必须考虑所有可能的颜色贡献的平均值与总和。这导致三种不同颜色因子的出现，这些因子就是用来表示三种过程（$q \to qg$，$g \to gg$以及 $g \to q\bar{q}$）相对强度的数值。这些因子的值可以直接归结到 SU（3）群的性质。夸克之间强相互作用的基本对称群是 $SU(3)_C$ 群，还是 SO（3）或者 U（1）群［实际上是指三个 U（1）群的乘积］？通过测量的颜色因子值，就可以找到答案。

凑巧的是，想要检验它们符合哪一种群，甚至都不需要直接

确定颜色因子的值：只需要知道它们之间的比率就可以了，而这对实验物理学家来讲，测量起来十分方便。三种颜色因子只能给出两个独立的比率。可以将这两个比率标记为比率 1 和比率 2。如果基本对称群是 SU（3）$_C$ 群，那么比率 1 的值应该是 9/4，比率 2 的值应该是 3/8。但如果强相互作用遵从 SO（3）对称性，则两个比率值都应该是 1。而基于 U（1）群的强相互作用所产生的比率值分别为：比率 1 为 0，比率 2 为 3，见表 4.4。对于最后一种情况，比率的差别主要是因为 U（1）群是一种阿贝尔群，因而图 4.35（d）所示的三胶子相互作用是被禁止的。但对于基于 SU（3）群的理论，三胶子相互作用是十分普遍的现象，因此四喷注事件中大约有 95%是 $q\bar{q}gg$ 过程，而只有 5%是 $q\bar{q}q\bar{q}$ 过程。

表 4.4 理论计算和实验测量得出的颜色因子。OPAL 的实验结果支持将 SU（3）群作为描述颜色的群

	理论			实验(OPAL)
	U（1）群	SO（3）群	SU（3）群	
比率 1	0	1	9/4=2.25	2.11±0.32
比率 2	3	1	3/8=0.375	0.4±0.17

在 LEP 加速器上有 4 台实验探测器，每一台都有一间房子那么大，并且每一台都是由世界上数百所著名大学或研究所组成的合作小组操纵的。按照字母顺序，第一台探测器是ALEPH，它的名字代表 LEP 物理仪器（Apparatus for LEP Physics），起这个名字本来的目的很可能是为了让它能在任何按字母排序的 LEP 探测器清单上都能占据有利位置。在这台机器上工作的一位研究生建议将名字改为“昂贵的大型硬件设备（A Large Expensive Piece of Hardware）”。名单上的下一台探测器是 DELPHI，或者叫轻子-光子-重子鉴定探测器（Detector with Lepton，Photon，Hadron Identification）。接下来的探测器让我们暂时摆脱了令人

痛苦的，有时互相纠缠的首字母缩写，它的名字叫做 L3；最后一台也是最为重要的一台机器就是 OPAL——“全功能 LEP 装置”(Omni Purpose Apparatus for LEP)，见图 4.36。4 台机器都使用了相似的技术来记录和识别粒子，但又各不相同，每一台都有自己特点。所有 4 台机器都用某种方法观察到了 Z 粒子的四喷射过程，以研究颜色群特征。

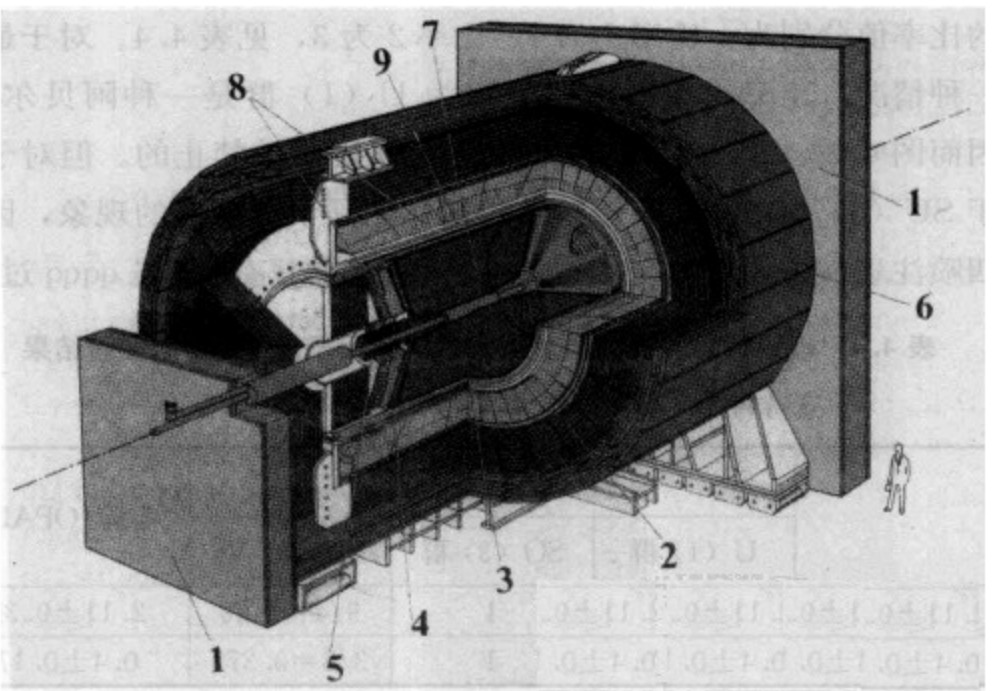


图 4.36 位于 CERN 的 OPAL 探测器的截断图。1. μ 子端盖；2. 强子量热器；3. 喷注室；4. 磁线圈；5. 前端探测器；6. μ 子管；7. 顶点探测器；8. 含铅玻璃；9. 预采样器和飞行时间探测器。图片来源：CERN

OPAL 合作小组在 1995 年发表了一项这类实验的结果，该实验包括了上百万次 Z 粒子衰变为强子的过程。OPAL 合作小组从上百万次事件中只发现了 21732 次四喷射事件，他们通过喷射角度对这些事件进行了一一的分析，见图 4.37。他们通过角度信息算出了颜色因子的比率值，分别是比率 1 为 2.11±0.32，比率 2 为 0.4±0.17；而基于三种颜色和 SU (3)$_C$ 群的 QCD 理论给出

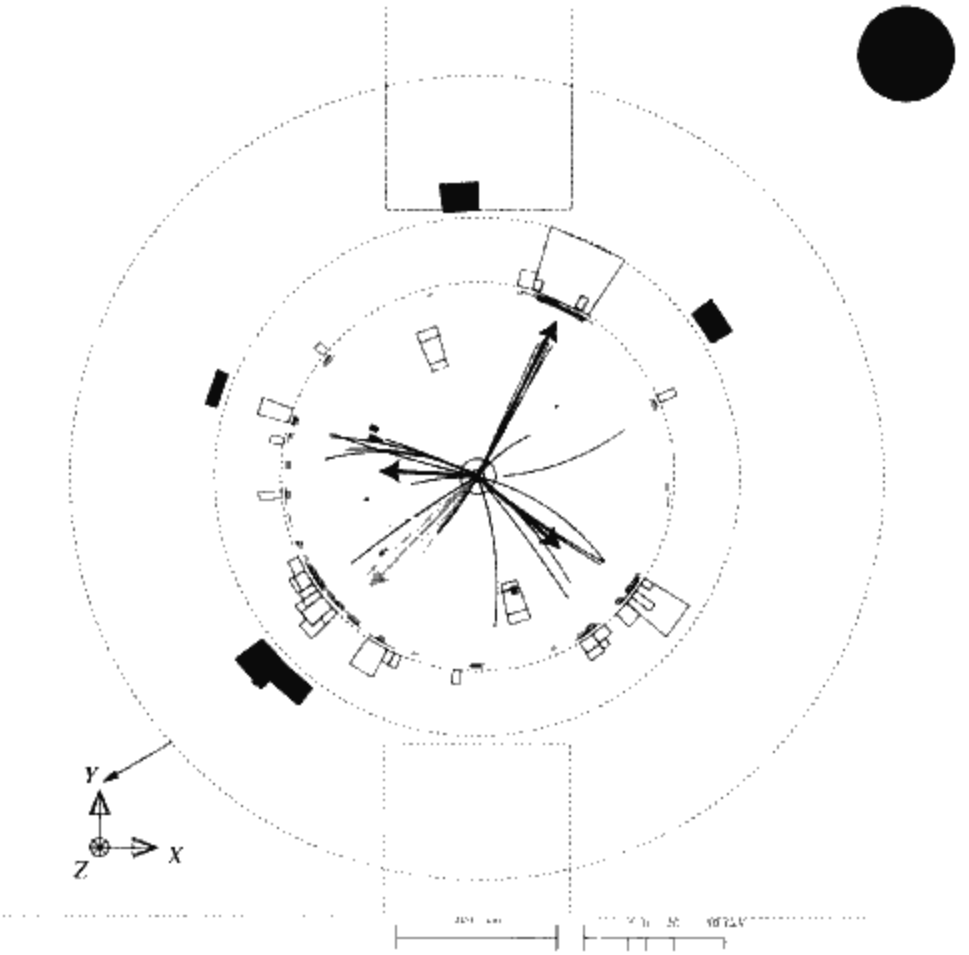


图 4.37 从 OPAL 探测器一端看过去，它记录下了在电子-正电子湮灭过程中的四喷注事件。箭头指示的是四条喷注。图片由斯特藩·克鲁斯制作。图片来源：CERN 和 OPAL 合作小组

的值分别是 2.25 和 0.375。

理论物理学家们对一长串候选群的颜色因子比率进行了计算。根据 OPAL 合作小组以及其他 LEP 小组的实验结果，许多群被直接排除掉了。但并不是所有候选群都能轻易地从名单中抹掉。例如，由于 OPAL 合作小组的实验结果误差太大，包含 4 种颜色的 SU（4）$_c$ 群也允许存在，至少在理论上如此。但是，就像上面讨论的电子-正电子湮灭过程，生成强子数量与 μ 子数量的

比率给出的结论那样，颜色的数量是 3 种，那么我们只剩下 3 种颜色群可供选择：SU（3）群，SO（3）群和 U（1）群。OPAL 合作小组的结果同 SO（3）群和 U（1）群都不符，因此它强烈地支持了 SU（3）群为颜色对称群的选择。

一套适合夸克的理论

尽管杨振宁和米尔斯在 20 世纪 50 年代以及后来的内山菱友已经构建起了现在被称作非阿贝尔规范理论的理论框架（QCD 理论就是其中之一），但仍留下了许多问题需要理论物理学家们解决。这种框架直接将基本对称性与物理中的相互作用联系起来，例如 OPAL 合作小组和其他实验者们的观测发现，实验上起作用的是 SU（3）群。但是，在用非阿贝尔规范理论对世界的基本物理进行描述之前，我们必须先回答两个关键问题。所有的量子规则能正确地包含进去吗？理论是可重正化的吗？这两个问题的答案都必须是“是”。

服从量子规则……

量子化就是强制服从量子规则的过程，在某个层面上，它表示将理论的主方程纳入费曼路径积分方法中，并导出费曼规则。但是这里却出现了一个问题。正是那种使规范理论非常成功的特点——规范不变性——同样也成为了问题的来源。它会引起一种重复累加，从而导致某些计算结果趋于无穷大。

这种重复累加的问题有点像在一个由传统家庭组成的小镇上进行民意调查，每个传统家庭都是妈妈、爸爸和几个孩子组成，而你可能会错误地认为小镇的家庭数就是小镇的人口数。由于每个家庭都包含几个成员，因此简单地计算人口数将会把家庭数估计过高，这就是重复计算问题的一个小例子。

路径积分方法要求包括理论中所有可能出现的场组态的贡献。问题就在于：对于规范场来讲，其中一些场的组态通过规范变换与另一些场有关。因此与只计算人数就会高估小镇上的家庭数的情况一样，用场组态进行计算就会导致重复计算（也称“多算”）。不仅如此，每个场组态可以通过规范变换与无穷多个其他场相关。因此直接运用费曼路径积分法则是不行的，你在最终的答案中会得到一个无穷大的数。

可以和民意调查的例子对比，就像城镇中有有限数量的家庭，但是每个家庭都有无穷多个孩子。因此如果只计算人数的话，就会得到一个无穷大的数，但如果只计算家庭就会得到合理的结果。同样，对规范理论进行量子化，就是找出一些方法，使每个家庭只选出一种场作代表，这样所有的场就自然按类别划分开了——规范变换联系着一个家庭内的所有场，但是却不能将不同家庭中的场关联起来。

每个场家庭挑选出一个成员的过程叫做规范固定（guage fixing），从许多可行的方案中决定怎样进行挑选叫做选择规范（choosing the guage）。从民意调查的例子来看，就像是决定通过计算成年男性的数量（而不是所有人的数量）的方法来决定家庭的个数一样。

施加了这种限制之后，每个用规范变换相联系的场家庭中只留下一种场，出现在路径积分中的无穷项最终并入一个无关大碍的整体常数项中，该常数项对画出至关重要的费曼图不产生任何影响。但是用这种方法限制路径积分会引发非常奇特的现象。当对规范固定后的主方程做费曼路径积分时，除了预期的费曼图之外，新加的限制还会额外导致费曼图的出现。新费曼图的出现说明，还存在一种新的神秘“粒子”。这些奇怪的新东西叫做鬼粒子（ghost）。

早在 20 世纪 60 年代早期，费曼首先意识到需要引入鬼粒子，

但两位来自列宁格勒（即圣彼得堡）的苏联研究人员，路德维希·法捷耶夫（Ludwig Faddeev，图 4.38）和他的学生维克多·波波夫（Victor Popov），真正用鬼粒子解决了规范多算的问题。因为这个原因，并且某种程度上也由于他们把这一点解释得最清楚明了，鬼粒子通常被称作法捷耶夫-波波夫鬼粒子。在得克萨斯，对量子引力论很感兴趣的布莱斯·德威特（Bryce deWitt）也指出了应当怎样利用鬼粒子，他的文章（比如在 1967 年发表的那篇）虽然很有预见力，但相对晦涩难懂一些。

那么鬼粒子到底是个什么东西？它们并不是物理实体，也就是说它们不是物理上真正存在的东西，不像夸克那样仅仅是不能观测到而已。鬼粒子在本质上是一种计算方法，是人造的产物，没有人会设计一次实验来寻找它们，也没有人想尝试这样做。它们没有自旋，但行为却像费米子——反鬼粒子和鬼粒子是不同的，因此它们的自旋统计是非物理的，这也是另一个不需要去寻找它们的理由。

鬼粒子只出现在含有规范玻色子圈的计算中。简单的树状费曼图中不会出现对应复杂鬼粒子图的额外部分。每一个规范玻色子圈都会有一个相应的鬼粒子圈图，见图 4.39。鬼粒子同规范玻色子之间耦合的细节依赖于规范的选择，但鬼粒子永远不会同电子和夸克这些费米子发生耦合。鬼粒子只是一种与理论中出现规范玻色子部分有关的性质，即使理论中不存在费米子，而只存在像胶子这样的规范玻色子，鬼粒子也必须存在。

引入一种非物理的粒子来赶走不希望出现的无穷大，这看起来有点像胡编乱造，好像是想从一种无关的理论中强行得出某些看起来有意义的结果。这种评价有点不公平。鬼粒子的出现是狭义相对论的要求，狭义相对论要求方程满足协变形式。这种要求就意味着，偶然会有非物理的规范玻色子贡献掺进来。这些东西必须被除去，但我们又没有办法在不破坏协变性的条件下做到这

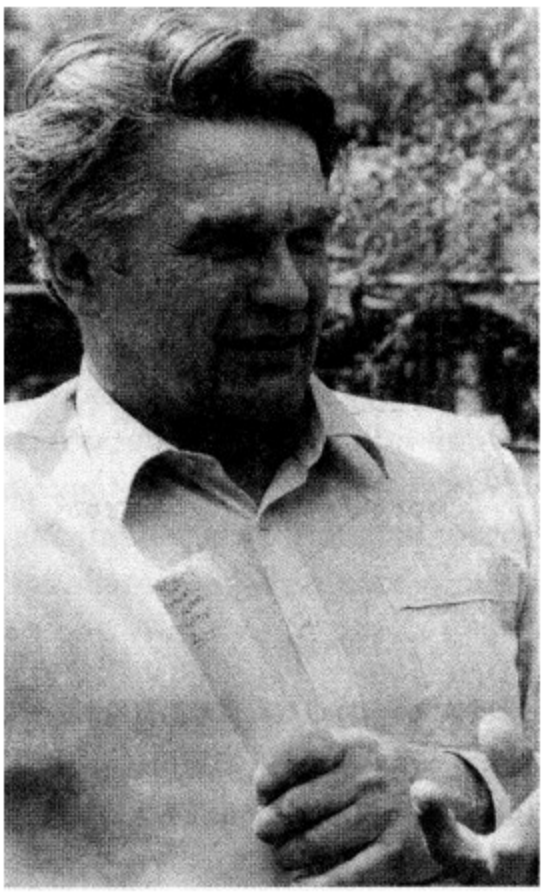


图 4.38 捉鬼人——路德维希·法捷耶夫，拍摄于 1986 年的阿鲁沙，当时他在一座俯瞰克里米亚海岸的杜布纳“度假村”参加会议。图片来源：路德维希·法捷耶夫/尤里·图曼诺夫（Yuri Tumanov）

一点。鬼粒子能在不破坏协变性的条件下，通过抵消多余部分解决这个问题。

我们还可以从另一个角度理解为什么需要鬼粒子。粒子自旋的方向相对于它的运动方向称为极化。自旋为+1 的光子自旋指向其运动的方向，也就是说极化是右手系的。极化为左手系的光子自旋与它的运动方向相反，其自旋值为－1。而平面极化光——从湖的表面反射出来的那种光，是两种极化光的混合产物。但是我们却没有自旋为 0 的情况。根据量子理论，在讨论自

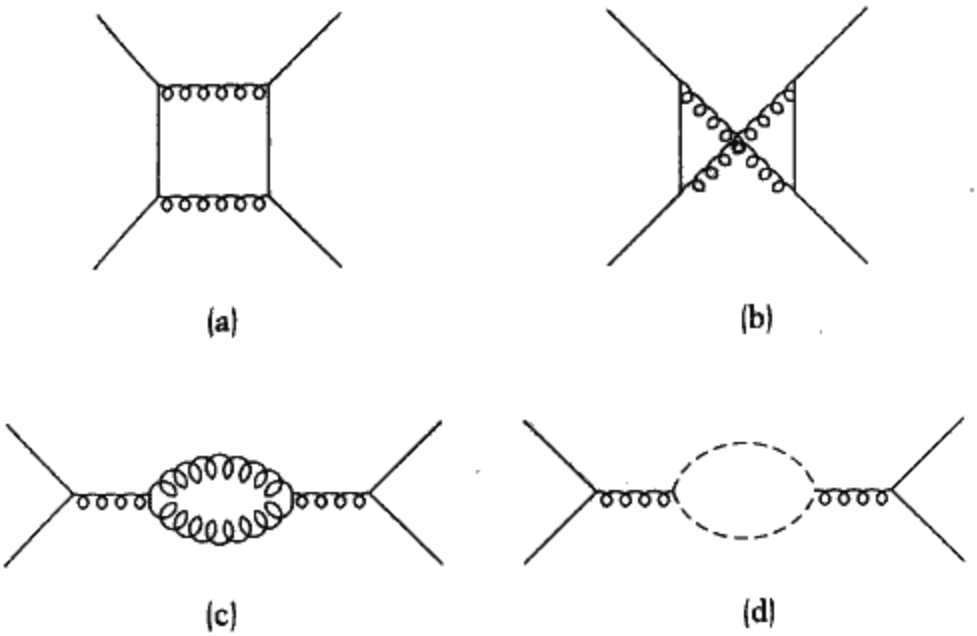


图 4.39 四个对 $q\bar{q}\to q\bar{q}$ 过程有贡献的费曼图——这些图都只含一个圈。图（d）表示一个鬼粒子圈，它同图（c）中的胶子圈相对应。如果没有图（d），其他三幅图就会产生无穷大的结果

旋为 1 的粒子时，其可以测量到的自旋值应为＋1、0、－1，这是一个很普通的性质。但是光子并不是自旋为 1 的普通粒子，因为它们没有质量，因此它们不存在自旋为 0 的情况。

物理上的胶子同样也没有质量，它同样也只有两个可能的自旋态；实验者们之所以不能观测到自然中存在左旋或右旋极化胶子，是因为大自然决定将它们隐藏起来，不允许人们观察到它们。问题在于在 QCD 理论中，带有封闭胶子圈的费曼图，仍包括自旋为 0 的胶子的贡献，这是洛仑兹协变性要求的结果。在最终答案中，这些非物理因素的贡献都应该被去掉。这项任务就要由鬼粒子完成，它抵消了由非物理胶子自旋产生的那部分贡献。它们在此过程中恢复了系统的幺正性——如果没有鬼粒子图，幺正性就会受到破坏。在特霍夫特和维尔特曼 1971 年做的关键工

作中，他们说明了：对于非阿贝尔规范理论来讲，鬼粒子图确实恢复了幺正性。法捷耶夫-波波夫方法的结果是一套与之一致的费曼图——包括树图和圈图。换句话说，非阿贝尔规范理论确实可以量子化。

那么为什么在QED理论中没有出现鬼粒子呢？规范固定问题在QED中同样也会出现，并且原因也是一样的，但是两者存在一点至关重要的差别。因为QED理论是一种阿贝尔理论，QED理论中的鬼粒子不会和光子耦合，也不会和任何其他东西发生耦合。因此鬼粒子的贡献可以分离出来，并安全地忽略掉。当然，将鬼粒子那一部分明确地表现出来也是可以的，如果有人非要这样做也未尝不可。鬼粒子在QED理论中不是一个问题，尽管理论物理学家们现在能够事后诸葛亮般地看出它是怎样“装配”在理论中的。QED理论在QCD理论和电弱相互作用理论之前被发现，可能是一件很幸运的事情，因为它不包含像鬼粒子这样复杂的内容。如果没有QED理论成功的指导和激励，先发现QCD理论以及一些额外的复杂内容就会艰难得多。

……击败无穷大

要想流芳千古有很多条路可以走。其中最有效的一个方式就是给科学杂志写一封短信，信中要包含一个特别重要的思想或方程。因为这个结果如此重要，所以有很多人会用到它，于是它就需要起一个名字，几乎可以确定的是，它的名字将会涉及提出它的那个人。约翰·沃德（John Ward）就是这样出名的。

在1949年，牛津大学的沃德在《物理快报》（*Physics Letters*）上发表了一篇半页长的论文，从此以后，该杂志场论专题那一章的副标题中就有了沃德的名字。当然，沃德并不是有意通过这种方法使自己出名。他关注的是戴森的一个猜测，即QED理论重正化过程中出现的三个无穷大常数中，有两个实际上是相同的。

沃德的短文证明了这一猜测是正确的。为了完成这项证明，他写出了一组以他名字命名的方程：沃德恒等式。在证明 QCD 理论和电弱相互作用理论可以重正化的过程中，沃德恒等式必不可少。

1957 年，爱荷华州立大学的高桥（Yasushi Takahashi）总结了沃德的工作。1971 年，法捷耶夫和波波夫的同事——莫斯科斯捷克洛夫数学院的安德烈·斯拉夫诺夫（Andrei Slavnov）和牛津大学的约翰·泰勒（John Taylor）把它推广到了非阿贝尔理论中。将沃德原来的理论同后来的进展结合起来，就构成了沃德恒等式。沃德恒等式是规范理论对称性的结果，实际上它表示的是格林函数之间的关系。格林函数，举一个例子来说，就是某些碰撞过程中“发生某件事情”的振幅。因此沃德恒等式保证了重正化方案中非常重要的消去作用。

值得信赖的规范理论模板——QED 理论是一个了解沃德恒等式意义的好地方。QED 理论中的基本相互作用涉及电子、正电子和光子，见图 4.40。图 4.40 的图（a）左边表示在电子-正电子-光子顶点图的基础上，所有可能的复杂费曼图的总和。图 4.40 的图（b）列出了其中的前两项。图 4.40 的图（a）右边的“费曼图方程”是有关传播子的，它描述了粒子从一点到另一点的传播。两个费米传播子之差除以两个传播子之积就等于左边的顶点图，其中两个传播子的动量差等于光子的动量。每个传播子本身都包含了所有费曼图的贡献。如图 4.40 的图（c）所示。我们可能不熟悉费曼图方程，但没有关系。重要的是顶点图——一个有关 3 个点的格林函数——同传播子之间存在联系，也就是说同有关两个点的格林函数存在联系。因此沃德恒等式联系了不同的格林函数。费曼图方程提醒我们，所有的费曼图都要包括进去——沃德恒等式不仅仅联系着具有相似复杂性的单个费曼图，而且对所有次数的微扰论都适用。

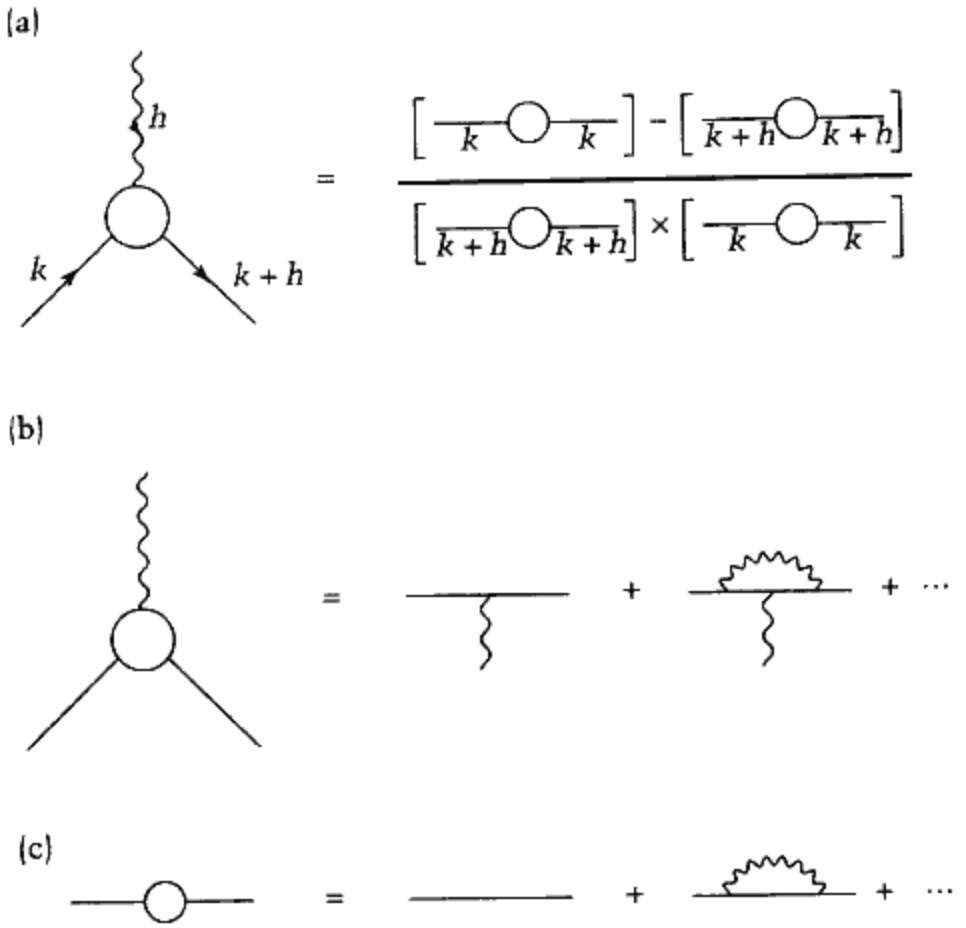


图 4.40 QED 理论中的沃德恒等式。(a) 顶点与传播子的关系。图 (a) 左边的顶点表示一个无穷多的费曼图序列，图 (b) 对其进行了详细说明。图 (a) 右边的每一个传播子都表示一个无穷多的费曼图序列，图 (c) 对其作了说明。符号 k 和 h 都表示动量

沃德恒等式是规范不变性的另一个副产品。回想一下我们上面着重提到的过程，即处理路径积分方法中固有的多算问题的过程。在某些阶段，我们必须选择一种特定的规范。一旦做了选择，必定会丧失规范不变性。但是在计算某些物理过程发生可能性的时候，规范的选择并不重要，只要方便就可以了，但必须要保留规范不变性。这两个明显矛盾的要求，导致了不同格林函数之间的关系——沃德恒等式。

我们该怎么做呢？沃德最初的论文中，说明了这些东西将怎

样帮我们解决即将遇到的问题——规范理论的重正化问题。QED理论的三个无穷大重正化常数中，有两个是相等的，这些常数联系裸量和观测量。由此，沃德证明了，最初需要三个常数的地方实际上只有两个常数是必需的。这个结果令人愉快，但实际上，其中隐藏着很深的物理意义。将QED理论用于其他带电物体之间的电磁相互作用——例如应用到μ子上——其实同将它用于电子和正电子之间的相互作用上是一样的。那两个常数相等表明，如果电子和μ子（以及其他种类的粒子）所带的没有经过重正化的电荷是相同的，那么经过重正化以后的电荷仍然是相同的：重正化是普适的，沃德恒等式证明了这一点。像电子和μ子这样不同种类的粒子具有同样的电荷，以及它们在QED理论中的耦合方式完全一样的说法，只有当沃德恒等式成立的时候才有意义。同样，在QCD理论中沃德恒等式也可以证明，所有服从规范性的性质——例如所有夸克无论带有什么颜色，都会以同样的基本方式同胶子耦合——在重正化后也同样成立。

但是在探索QCD理论方面，用沃德恒等式证明理论可以重正化是非常重要的。可重正化的证明依赖于搞清费曼图之间复杂的抵消关系。例如在QCD理论中，有来自于胶子的非物理贡献，必须用鬼粒子的贡献将其抵消，这样理论才能给出有意义的预测结果。沃德恒等式可以有效实现这些抵消。特霍夫特正是通过沃德恒等式，证明了整个杨-米尔斯理论系列的可重正性，因此也证明了QCD理论的可重正性，他的这一工作为自己赢得了诺贝尔奖。这对一个研究生来说，已经很不错了。

QCD理论诞生

科学家们到底应该在什么时候庆祝QCD理论的生日呢？QCD理论到底多大年纪了？QCD理论是一套杨－米尔斯理论，其中的基本粒子是夸克，规范群是颜色SU（3）群，传递力的粒

子（胶子）是一个“八人小组”，并且每种胶子都带有颜色。韩武荣和南部阳一郎早在1965年就提出了以上内容，尽管在他们的模型中，夸克还带有整数倍的基本电荷。他们还注意到，可观测的强子整体上是颜色中性的，换句话说它们是颜色单态。这就是QCD理论的诞生吗？答案是否定的。也就是说，通常人们并不承认他们的工作就是QCD理论的开端。他们的思想好像已经超越了他们所在的时代，那些最终能编织到一起组成QCD理论的线还没有纺成呢。

这些线中的一条（这些线并不是完全独立的）与杨-米尔斯理论有关。杨-米尔斯理论揭示了太多东西。大约在1967年，格拉肖、温伯格和萨拉姆开始构建关于弱力的杨-米尔斯理论。但没有人知道，该理论或者其他的杨-米尔斯理论是否可以重正化。1971年，特霍夫特和维尔特曼证明了它们都是可重正化的。直到1971年前，杨-米尔斯理论还被人们丢弃在布满灰尘的角落里，但在1971年，人们突然再次对它产生了兴趣。格拉肖、温伯格和萨拉姆的电弱理论复苏了，人们很自然地想到用杨-米尔斯理论也构建一套描述强相互作用的理论。

另一条线就是令人无法抗拒的新量子数——颜色量子数。从1971年末开始，盖尔曼同德国理论物理学家哈拉德·弗里奇（Harald Fritzsch）以及来自普林斯顿的威廉·巴丁（William Bardeen）共同在CERN工作了一段时间，见图4.41和图4.42。巴丁尤其善于解决中性π子衰变为两个光子的复杂问题，由于π子由带分数电荷的夸克所组成，因此当时衰变率的计算值只有测量值的1/9。

1972年，在一篇首先作为CERN的预印本出现的论文中，巴丁、弗里奇和盖尔曼证明了只有在每个夸克具有3种颜色的情况下，才能够得出正确的衰变率计算值。在同一篇论文中，他们还证明了由于颜色的存在，电子-正电子碰撞所产生的强子与μ

图 4.41 哈拉德·弗里奇（左）和默里·盖尔曼在一次柏林的会议中对着镜头微笑。图片来源：哈拉德·弗里奇

子的比率要乘以一个值为 3 的因子，这在前面已经提到过了。这两个发现至今仍是人们最经常引用的证明颜色存在的例子，当然还存在更多的证据。也是在 1972 年，巴黎大学的理论物理学家克劳德（Claude Bouchiat）、约翰·里尔普洛斯（John Iliopoulos）和菲利普·迈耶（Philippe Meyer）证明，颜色属性对于保持包含夸克和轻子的电弱理论的可重正化性十分重要。因为这种抵消过程十分精美，该结果——戴维·格罗斯（David Gross）和罗曼·贾基夫（Roman Jackiw）也导出了这个结果——对理论物理学家有着独特的吸引力。

还是在 1972 年，盖尔曼代表他自己和弗里奇在国家加速器实验室（这是费米实验室过去的名字）举行的第十六届国际高能物理会议上宣读了一篇论文。他们的论文提出了一个强相互作用模型，模型中包含了带颜色的夸克和一个不带颜色的胶子。所以

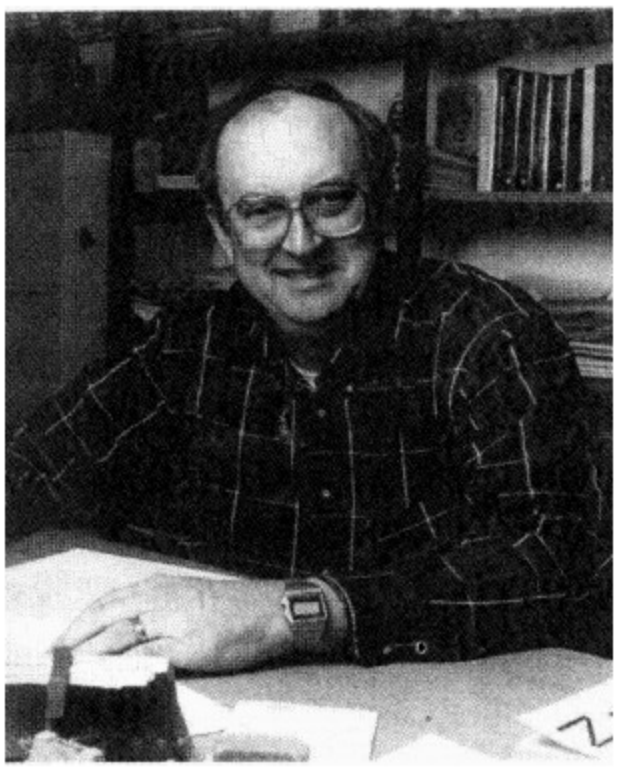


图4.42 比尔·巴丁1996年摄于他在费米实验室的办公室中。巴丁的父亲约翰·巴丁是唯一一个两次获得诺贝尔物理学奖的人：第一次是因为发现了半导体的晶体管效应，第二次是因为提出了关于超导的巴丁-库珀-施里弗理论（常被称为BCS理论）。图片来源：费米实验室

这并不是我们今天所知道的QCD理论。但是他们的论文中还是有很多重要的概念。论文中称“真实的粒子必须是颜色SU（3）群的单态”，这是对他们1972年那篇尝试性的预印本以及韩武荣和南部阳一郎工作的再次强调。弗里奇和盖尔曼还提到了夸克和胶子都应该带有颜色的思想。他们称：“例如，它们能够形成一个中性矢量场的颜色八重态，矢量场满足杨-米尔斯方程。”这听起来很像QCD理论。但是对这种方法的潜在优点进行了两次评估之后，他们还是选择了单胶子方案。那么这篇论文能代表QCD理论的诞生吗？有一些理论物理学家同意这种观点。

第三条线提供了 QCD 理论诞生的另一个候选方案。1969 年，SLAC 的实验者们证明，质子看起来包含一些点状的散射中心。该发现给理论物理学家们带来了严峻的挑战，这是一种被叫做“比约肯标度无关性”（Bjorken scaling）的现象，我们稍后再讨论它。它对普林斯顿的戴维·格罗斯和他的学生弗兰克·威尔茨克（Frank Wilczek）以及哈佛大学的研究生戴维·波利茨(David Politzer）研究重新流行的杨-米尔斯理论在高能情况下的行为是一种鼓舞（图 4.43 至图 4.45）。两个小组都指出，杨-米尔斯理论的相互作用强度，随着能量的增加会逐渐减小，直至消失。格罗斯和威尔茨克特别注意到，运用杨-米尔斯理论，“一个人能够构建许多有趣的强相互作用模型”。在得知出现颜色量子数的证据之后，格罗斯和威尔茨克根据巴丁、弗里奇和盖尔曼在 1972 年关于 π 子衰变和电子-正电子湮灭的工作，描述了“一个特别吸引人的模型”——这听起来仍十分不确定。在该模型中，颜色 SU（3）群容纳了强相互作用。具有同样味道但不同颜色的夸克通过电中性的“矢量介子”（即胶子）混合起来。现在，这听起来很像 QCD 理论了。格罗斯先写下了杨-米尔斯理论的主方程，然后确定了他们“吸引人的模型”中的 QCD 颜色群对称性，他声称他和威尔茨克是最先发表 QCD 主方程的人。一些人肯定会争论说：一旦把杨-米尔斯理论作为框架，并且确定所用的群是 SU（3）群，写出主方程是水到渠成的事情。我们先把这些放到一边，用现代夸克和胶子的语言来讲，格罗斯和威尔茨克以及独立得到该结果的波利茨发现的是：当夸克离得越近，夸克之间由胶子所传递的力越弱。在很高的能量下，夸克的行为越来越像自由粒子，就好像强力松开了束缚一样。这种性质被称作渐近自由(asymptotic freedom)，它被公认为是 QCD 理论的基础。格罗斯和威尔茨克不久又提交了两篇很长的论文，这两篇论文补充介绍了渐近自由及其相关结论，并介绍了颜色 SU（3）群的普遍情

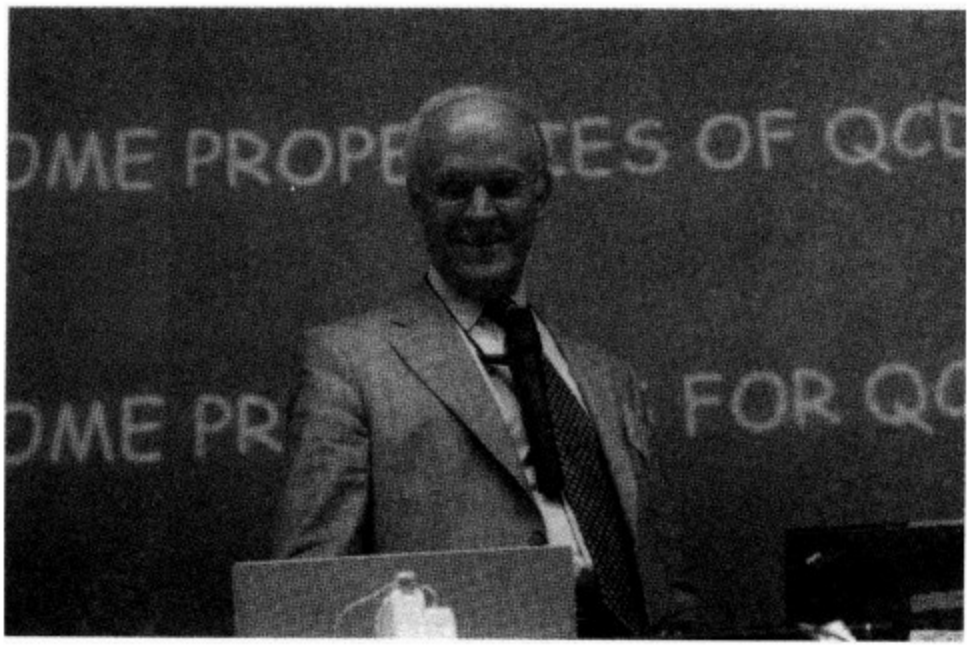


图 4.43 戴维·格罗斯在会议讲坛上，摄于 2003 年。图片来源：戴维·格罗斯

况，他们明显已走在了发现 QCD 理论的前列。

一些教科书还将由弗里奇、盖尔曼和海因里奇·鲁特威勒（Heinrich Leutwyler，图 4.46）所写的另一篇论文列入 QCD 理论的一般参考书。这篇于 1973 年完成的文章，标题为“颜色八重态胶子图像的优点”，是紧随着格罗斯、威尔茨克和波利茨的论文发表的，文章的标题说明格罗斯等人的工作是一次进步。在这篇文章中，盖尔曼和他的同事最终抛弃了单胶子的观点，转而将注意力集中到了正确的 QCD 理论上来。现在，所有关键的要素都具备了：主方程；格罗斯、威尔茨克和波利茨提出的渐近自由；可观测到的粒子都是颜色单态的思想；杨-米尔斯理论；正确的颜色对称性 SU（3）群。然而这里一点也没有发现新事物时的激动心情，更多的则是这样一种感觉：你同时看着一辆令人着迷的新跑车和一辆曾经深深喜爱过但已经被撞毁而只能扔进垃圾

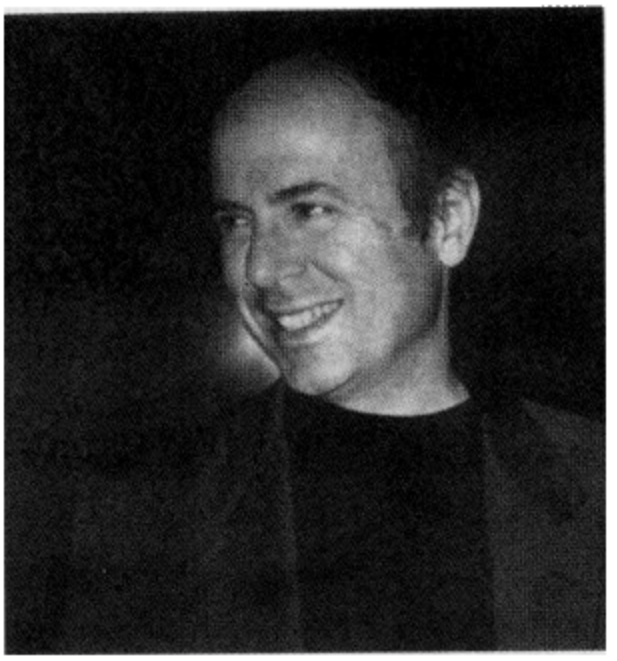


图 4.44　弗兰克·威尔茨克参加费城举行的美国物理协会会议，摄于 2003 年。图片来源：伊丽莎白·戴维妮

堆的旧汽车。公正地讲，单胶子模型本质上是盖尔曼在研究夸克时用的一种工具，他在 20 年之后将其形容为一个“受遗弃”的模型。

就人们公认的想法来看，QCD 理论是在 1973 年底诞生的，尽管那时它还没有名字，也没有突然出现的感觉。当理论发展了一年左右后，人们开始逐渐接受它了。

“渐近”是一个数学术语，用来表示一个量接近但永远无法到达某个特定值，这个特定值通常是零或者无穷大。该语境中的“自由”是指不受相互作用影响。渐近自由——力的强度随着能量增加而减小的现象——是 QCD 理论的核心内容，它对于利用 QCD 理论进行计算，对于理解我们为什么看不到自由夸克和胶子都有重要的意义。实际上，就像格罗斯、威尔茨克和波利茨所展示的那样，它是非阿贝尔规范理论的普遍性质。好像要更充分地

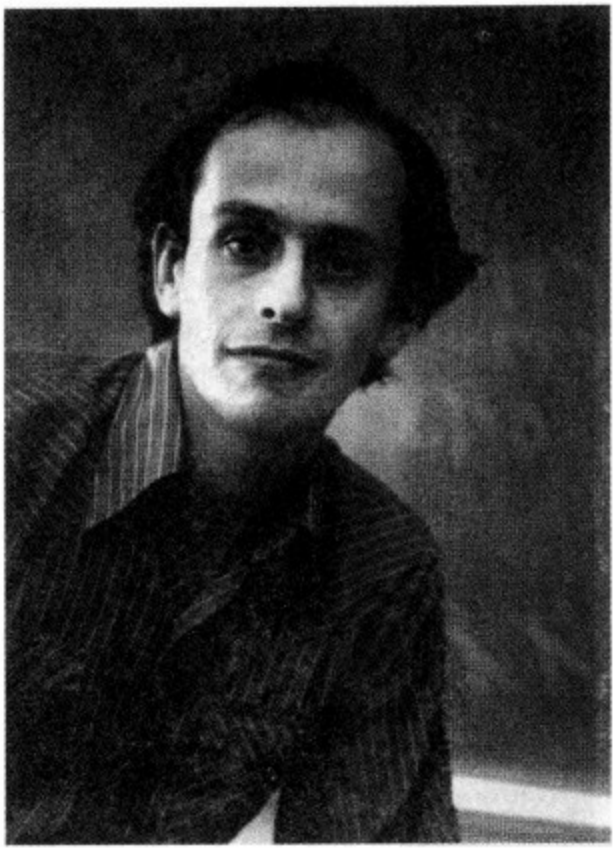


图 4.45　戴维·波利茨摄于 1979 年。波利茨与格罗斯和威尔茨克分别独立发现了渐近自由的关键性质。图片来源：加利福尼亚理工学院

说明这一点，格罗斯和波利茨的论文导师西德尼·科尔曼（Sidney Coleman）还证明了：如果可重正化的场理论不具有非阿贝尔性，就不会具有渐近自由这种性质。因为标度无关性要求渐近自由的存在，一套能解释 1969 年 SLAC 实验结果的理论，必须是一套非阿贝尔规范理论。

渐近自由的发现有其自身的数学基础，重正化群是其中的关键部分。它将理论中“短程”和“长程”概念之间固有的任意性转变为像质量、相互作用强度这样的依赖于能量的关键参数，这一点在第五章中还要讲到。瑞士物理学家恩斯特·斯特科伯格

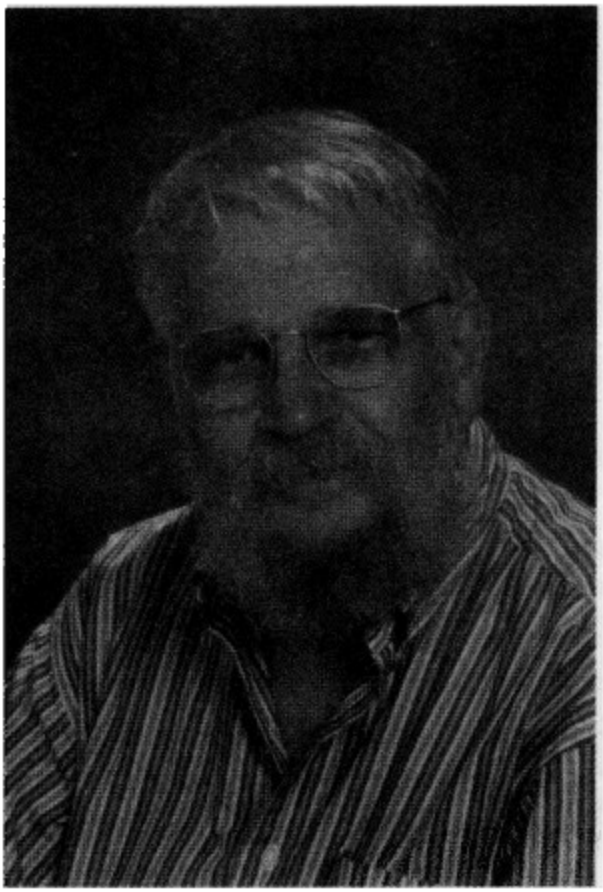


图 4.46　海因里奇·鲁特威勒在波恩，摄于 2003 年；弗里奇、盖尔曼和鲁特威勒在 1973 年发表的论文中讨论了 QCD 理论的结构。图片来源：海因里奇·鲁特威勒

（Ernst Stückelberg）和安德列·彼德曼（André Peterman）于 1953 年首次发现了重正化群。但盖尔曼和伊利诺伊大学的弗朗西斯·洛（Francis Low）也独立发现了重正化群，并将它用到 QED 理论的研究中。用现在的说法，他们证明了 QED 理论不存在渐进自由。实际上，QED 理论的情况同 QCD 理论完全相反——在能量越高的情况下相互作用的强度会增加。

意大利理论物理学家吉尔吉奥·帕里西（Giorgio Parisi）提出，点状物的散射行为要求强力在能量升高时减弱。在他的指引

下，格罗斯、威尔茨克和波利茨用重正化群研究了杨-米尔斯理论。他们发现，在高能情况下，耦合参数被限制在某个固定值上，而这个值就是零——也就是存在渐近自由。富有戏剧性的是，在他们早期的计算中由于弄错了一个负号，格罗斯和威尔茨克本来还以为他们证出了完全相反的结论。

在柯蒂斯·卡兰（Curtis Callan）（图 4.47）、科特·西曼吉克（Kurt Symanzik）、肯尼斯·威尔逊（Kenneth Wilson）、温伯格、特霍夫特以及其他一些人的努力之下，重正化群和渐近自由相互交织的发展，形成了一座新的里程碑。1970 年，加州理工学院的卡兰和 DESY 的西曼吉克各自独立地推导出了一套重正化群方程，这些方程能够处理有质量的夸克。他们又重新发现了列夫·欧凡什安尼科夫（Lev Ovsiannikov）曾在 1956 年得到的结果。在 20 世纪 50 年代，谢科夫（Dmitrij Shirkov）、尼古拉·波

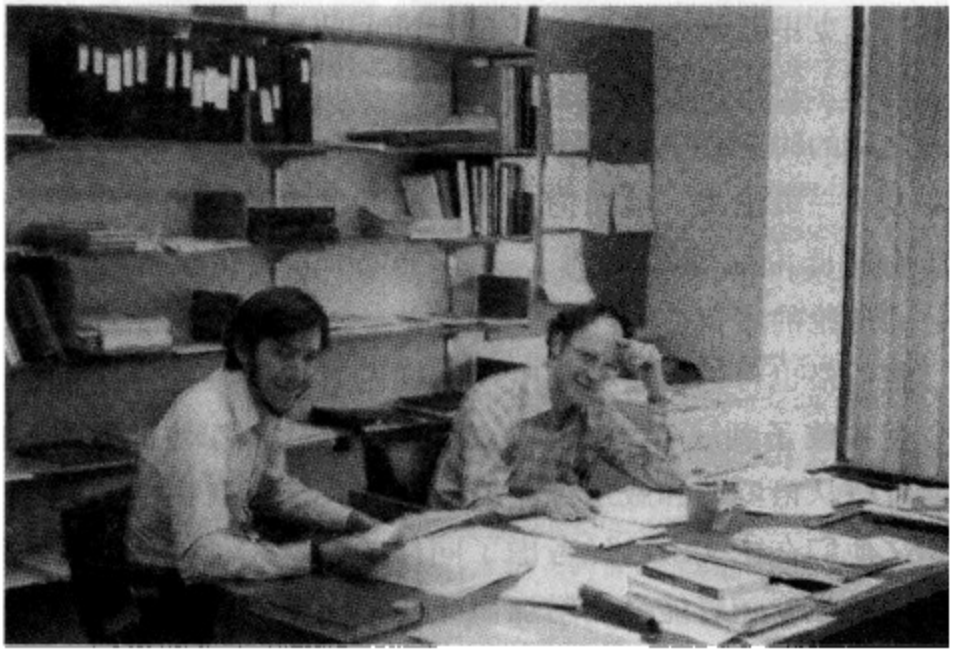


图 4.47　渐近自由的曙光：柯蒂斯·卡兰（左）和戴维·格罗斯，摄于 1972 年。图片来源：戴维·格罗斯

戈留波夫（Nicolai Bogoliubov）等一些苏联物理学家，在发展重正化群思想的方面做出了大量的工作。逐渐被人们所熟知的卡兰-西曼吉克方程成为人们发现渐近自由的起点。威尔逊最开始是盖尔曼在研究强相互作用方面的学生，他将重正化群的思想应用于凝聚态物理，更确切地说是所谓的“临界现象”（critical phenomena）中。开展于20世纪70年代早期的这项工作，不仅使他获得了1982年的诺贝尔物理学奖，还确立了重正化群的地位，使它成为在大范围内解决物理学问题的强有力工具。但事实上，渐近自由的开端要追溯到1965年，那时两位苏联物理学家弗拉德米尔·范尼亚金（Vladimir Vanyashin）和米哈依·特仑特夫（Mikehail Terentev）设计了一套杨-米尔斯模型理论，他们在计算过程中注意到一种他们不愿看到的情况，即相互作用强度随距离缩短而减弱，与电动力学中距离越短相互作用越强的情形相抵触。这就是渐近自由的信号。四年之后，另一位苏联物理学家伊奥斯夫·克里普洛维奇（Iosif Khriplovich）也推导出了相似的结果。特霍夫特也同样如此，在1972年的马赛会议上，西曼吉克演讲之后的讨论会中，特霍夫特提到了他的工作。1973年，温伯格发表了一篇论文，他在论文中将渐近自由和弱相互作用及强相互作用的群论原理结合到了一起，拉开了现在被称为“标准模型”的理论的序幕。

温伯格也讨论了这种刚出现的理论直接面临的一个问题：自由夸克和胶子到底存不存在？理论物理学家们现在所理解的答案，可以从格罗斯和威尔茨克第一篇有关渐近自由的长篇论文中找到，他们在论文中已经考虑了同样的问题。他们指出，如果相互作用的强度随着能量的增加而减弱，在低能状态下相互作用强度会相应地增加。他们的想法是，在低能状态下，夸克之间的耦合如此之强，以至于我们永远也看不到自由夸克——夸克们被限制住了。这就是夸克禁闭的概念，即夸克和胶子总是被封闭在复

合粒子中。实验物理学家们通过他们的探测器只能观测到颜色中性的粒子。由于存在渐近自由，夸克禁闭是一种相当合理的想法，但这种想法永远也不可能被严格证明，我们会在第七章进一步探讨这个问题。这种相互作用强度在低能（也可以等价地说是“长距离”）条件下激增的事实，另外还为胶子传递过程的短程性提供了自然的解释。反之，像光子和引力子这样无质量的携带力的粒子，是与无限范围内的相互作用相联系的。

渐近自由是QCD理论最重要的特征之一，我们可以说，如果没有这个发现，QCD理论就是不完整的。人们为解释令人吃惊的实验结果进行了不断的探索，从根本上推动了这次发现的完成。

夸克一瞥

用红色灯泡拼成的“比萨店”招牌不如绿色的好辨认，而绿色的又不如蓝色的容易辨认。至少理论上应该是这样的。现实情况会稍微复杂一些，因为首先，人的眼睛就不是一部完美的光学仪器。但波长的长短和仪器（可能是望远镜或者人的眼睛）成像是否清晰之间实际上存在着联系。

光学上有一条黄金定律，波长越短成像就越清晰，什么东西都不例外。因为波长越短，光就越不容易因固有的波动性而产生衍射，而衍射会限制分辨率的提高。用过显微镜的人都知道这些内容。他们经常用到的光学显微镜观察细节的能力——换句话说，也就是显微镜的分辨率——由于光的波动性而存在一个理论极限，大约是几百纳米。要想看到更小的细节，显微镜就应该使用更短的波长。这就是为什么要发明电子显微镜。这是对物质“粒子”波动本性的直接应用：电子显微镜中的电子通过10～200千伏的电压进行加速，其波长小到几十纳米，精度足以看到大分子或聚合分子。

因此，想要看到更小细节，其方法就是利用非常短的波长。质子的直径大约是 1 飞米（10^{-15}米），要想看到这么小的东西，或者甚至想要看到比它还小的东西，就需要一台十分大的电子显微镜。德国北部城市汉堡就坐落着一台（从效果上讲）世界上最大的电子显微镜。

HERA——世界上最大的电子显微镜

1995 年，HERA（the Hadron-Electron Ring Accelerator，位于德国 DESY 的强子-电子环状加速器）的 H1 合作小组成功地证明了，轻夸克（指 u 夸克、d 夸克和 s 夸克）的半径小于 2.6×10^{-18}米。这还不能说明夸克是没有大小的，当然那也是很有可能的。该结果只是表示夸克非常非常的小。“电子显微镜”HERA 使得这次测量成为可能，它将探测的距离降低到了约为原子大小的一亿分之一。

HERA（图 4.48）在 1992 年投入使用，它开创了一个加速器的新时代。它是一台对撞机，但并不是将两种相同的粒子对撞——既不像在 LHC 的质子-质子对撞过程那样，也不像在 LEP 和 Tevatron 上那样实现粒子-反粒子对的对撞，而是将两种完全不一样的粒子——电子和质子实现对撞。HERA 由一个 6.3 千米长的圆形管道组成，管道中包含了两个独立的加速环。上面的环利用强大的超导磁铁，引导质子沿环的一个方向前进。它的正下方是另一个环，用低功率的普通磁铁引导电子（或正电子）向相反的方向运动。两个环在两个位置相交，粒子束可以在那里发生对撞。

在这些相交处放置着两台探测器 H1 和 ZEUS［在希腊神话中，HERA（赫拉）是 ZEUS（宙斯）的妻子］，用来观测电子-质子对撞。这些探测器要经受特殊的挑战，因为尽管 HERA 是一台对撞机，但质子所携带的动量要比电子大得多。这很容易造

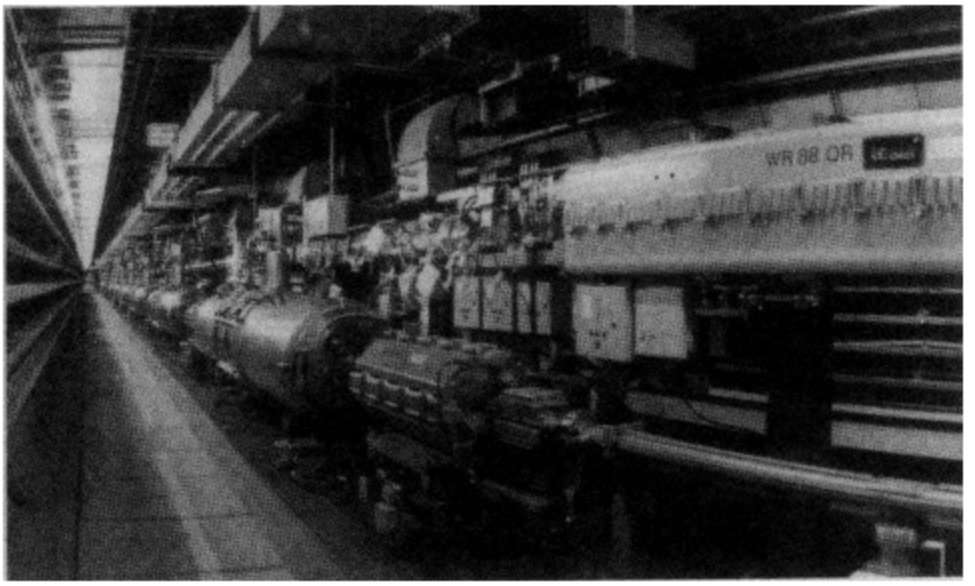


图 4.48　在 6.3 千米长的 HERA 隧道下拍摄的照片。图片来源：德国汉堡的 DESY

成碰撞产物向着质子运动方向的一个顶角很小的圆锥范围内喷射。质子碎裂后的碎片，很容易继续沿着质子束运动的管道前进，而无法被探测到，见图 4.49。

通过将 920 吉电子伏的质子和 27 吉电子伏的电子对撞，可得到的总有效碰撞能量为 318 吉电子伏，因此 HERA 所观测到的核子内部的情况要比其他任何机器清楚 100 倍以上。HERA 耗费了 6.5 亿美元的资金，换来了大约为质子直径千分之一大小的分辨率；它只耗费了商业电子显微镜成本的 1000 倍，却提供了比商业电子显微镜分辨率高 10^{10} 倍的放大倍数。这真是物超所值了！

那么 HERA 到底看到了什么？一个迎面而来的电子被带电的质子组分弹开，射入了用来记录它能量和散射角度的探测器中。这是一种电磁相互作用，因此是通过光子传递的。该过程可以用最简单的费曼图清楚地表示出来，见图 4.50，图中入射电子放出了一个光子，正是这个光子插入了质子内部。从效果上看，目标已经被散射电子发出的光子探测到了。具有最大动能的光子

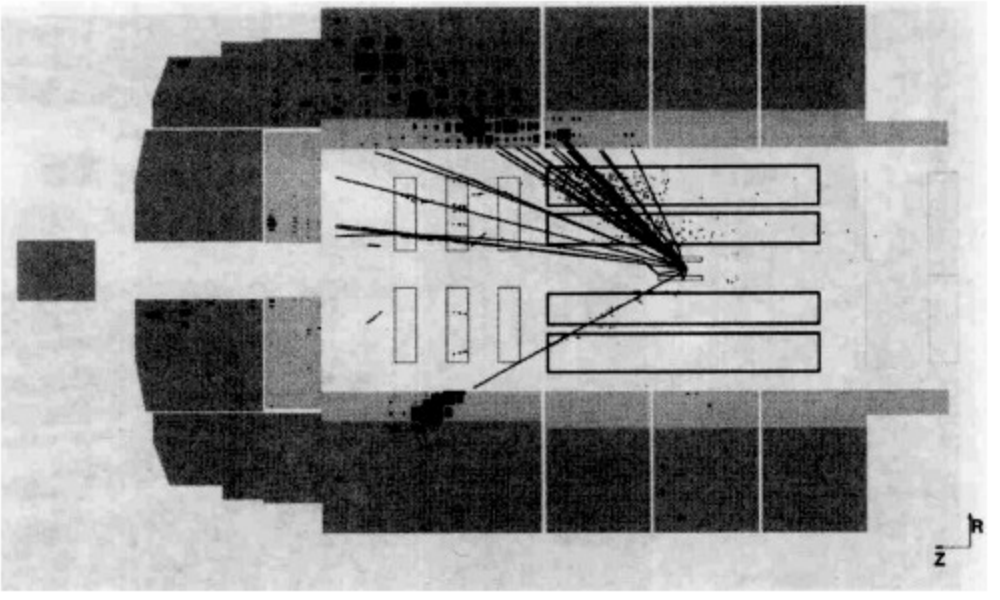


图 4.49　由 H1 探测器俘获到的电子-正电子碰撞事件。图片来源：德国汉堡的 DESY。图片来源：戴维·格罗斯

能够产生最猛烈的碰撞，它具有最短的波长，因此能够显示出最清晰的核子结构图像。这是一种深度的非弹性散射，由于是非弹性散射，电子在撞击质子的时候会损失能量，媒介光子所具有的波长足够短，因此它能够刺入质子的深度也很深。举例说明，电子在质子上的深度非弹性散射可以表示为 $e+p\rightarrow e'+X$，其中 X 表示将质子炸成碎片之后所得到的产物，e'表示被散射的电子。深度非弹性散射是研究核子结构的一种主要工具。

在粒子物理中，通常用两个关键的数字描述这种散射。其中一个是电子损失的能量，也就是入射电子和出射电子的能量之差，因此它被称为能量迁移。另一个是由光子传递给核子的四维动量（平方的相反数），这在第三章中已经讨论过了。这种“碰撞的强度”可以通过入射电子和出射电子的能量来衡量，也可以通过电子被散射的角度来衡量。那个光子是一个普通的真实光子吗？如果是的话，我们就不需要用到第二个数值了，因为真实光子是没有质量的，因此真实光子四维动量的平方总是零。但那个

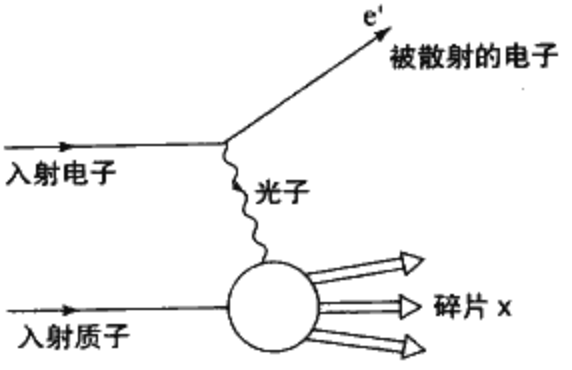


图 4.50 电子-质子的深度非弹性散射。入射电子会放出光子，光子将质子撞成碎片 x

插入质子心脏中的光子是虚光子，它的有效质量是通过这种动量迁移给出的。动量迁移越大，虚光子的波长越短，实验中所能分辨的细节就越精细。

HERA 的研究者们以及进行其他深度非弹性散射实验的实验者们，会测量电子-质子碰撞的反应截面。也就是说他们会寻找和测量散射电子，而忽略被击碎的质子所留下的残骸。只要入射粒子愿意，它就能够通过光子看到另一个与它自己相像的带电物体，这个物体具有一定的能量和动量范围。这是因为，尽管质子的能量和动量大家都知道，但组成质子的带电散射体及其动量和能量都是不确定的。实际上，位于质子中散射中心的能量和动量相对于质子本身的能量和动量的分配，可以通过某个物理量给出，物理学家们将其称为结构函数。他们能够通过测量散射截面得到这个值。深度非弹性散射揭示了这种带电的点状散射中心——夸克——在核子中的分布，随着 HERA 的启动，实验物理学家们正在进行的这个实验马上会达到一个新的顶峰。

回旋加速器的问题

全世界散布着许多回旋加速器，有一些被用作研究，有一些

用于医疗——中子在摧毁癌肿瘤细胞方面的效果比 X 射线要好得多，质子在传递消灭肿瘤细胞的辐射，但同时又不破坏附近的组织方面表现得也比 X 射线要好。回旋加速器还可以用来制造医疗上使用的放射性同位素，还能用于制造高级微电子芯片。

但是如果利用回旋加速器生产高能粒子束，来进行粒子物理方面的研究，就会出现一个问题：当粒子接近相对论要求的速度时，时机的选择就显得特别重要，一旦选择不慎，整个机器的功能将会彻底崩溃。其解决方案之前已经提到过，就是要稳定相位，这导致了同步回旋加速器和回旋加速器的诞生。

但是还存在第三种利用相位稳定性的粒子加速器，那就是直线加速器，或称作 linac，粒子在直线加速器中可以在一条直线上进行加速。一条笔直的粒子束（而不再是环形的粒子束）会沿着带电圆柱的轴线前进。加速发生在圆柱体之间的空隙处，而在每个单独圆柱内部运动的时候，粒子的速度不改变。从加速器的一端沿着直线开始数，排在奇数位置的圆柱都是用导线连在一起的，偶数位置的圆柱也同样用导线连接在一起，见图 4.51。只要粒子在圆柱中运动的时间与电压变换的时间相符，从圆柱中出来的粒子就能及时赶上加速电压。

相位稳定性是用圆柱长度控制的，在能量较高的一端圆柱的长度也较长。如果粒子到达圆柱的时间太晚，它就会受到一次更强的冲击，这样，到达下一个空隙的时间就会稍微提前一些。如果粒子到得太早了，它受到的冲击就会比较小，到达下一个地方也会比平时稍微晚一些。尽管可以将粒子连续地导入加速器的起始端，但只有那些恰好赶上周期的粒子才能够被加速——因此这台机器输出的也是一系列由加速粒子形成的脉冲。从本质上讲，有一列电磁波在加速器的整个长度上传递，电子被它裹挟着，在电磁波上“冲浪”。

直线加速器与回旋加速器相比有一些重要的优点。由于粒子束

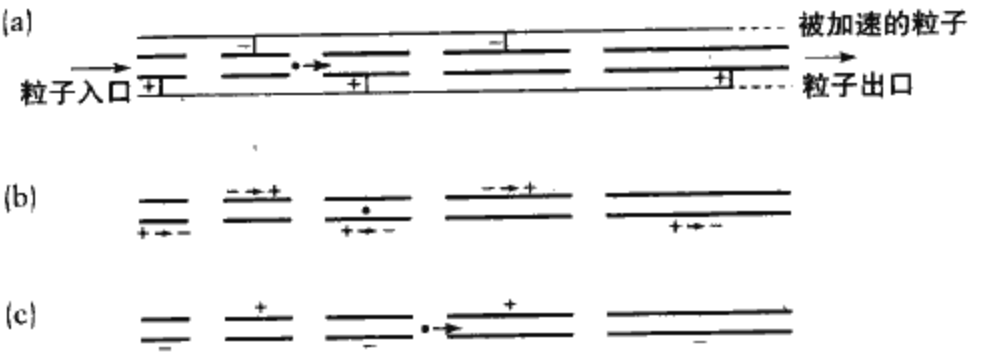


图 4.51 直线加速器的示意图。在图（a）中，导线将圆柱连接起来。一个带负电荷的粒子正朝第三个圆柱加速运动。图（b）：当粒子进入第三个圆柱后，圆柱上所加的电压发生反转，用从 + 到 − 的箭头表示。图（c）：因此，当粒子再次离开第三个圆柱，它会继续向着带正电的第四个圆柱加速前进

是直线的，就不需要昂贵的偏转磁铁。加速电压的频率是常数，直线加速器中的粒子束同回旋加速器中的粒子束相比，能更好地汇聚。其他的优点有：因为粒子是直线传播的，因此不会发出同步辐射，这是像电子这样质量很轻的粒子特别需要注意的一个问题。

最早的一台大型电子直线加速器是位于斯坦福那台具有 1 吉电子伏能量的 Mark Ⅲ。罗伯特·霍夫施塔德（Robert Hofstadter）所带领的小组，在 1956 年正是利用这台机器研究了电子和质子之间的弹性散射过程，他也因此获得了 1961 年的诺贝尔物理学奖（图 4.52）。霍夫施塔德和他的同事测量了散射截面同电子散射角之间的依赖关系，最终证实了质子并不是一个点状物体，而是具有确定大小的粒子。入射电子将质子看做一个具有可测半径的带电球体，霍夫施塔德和他的同事测定质子的半径为 0.74×10^{-15} 米，或者说是 0.74 飞米。霍夫施塔德的结果是人类第一次得到核子具有大小的直接证据。

图 4.52　罗伯特·霍夫施塔德，他进行了电子同质子的散射实验，证明质子并不是一种点状粒子，而是具有有限的大小。霍夫施塔德获得了 1961 年的诺贝尔物理学奖。图片来源：斯坦福大学，美国物理学会艾米利奥·西格雷可视档案馆，W. F. Meggers 诺贝尔桂冠陈列室

质子中的部分子

Mark Ⅲ加速器反过来又刺激了 3 千米长的 SLAC 加速器的出现，有时它也被称为“两英里直线加速器”。这台加速器于 1966 年开始使用，最一开始可以将能量为 17 吉电子伏的电子打到固定靶上。在 20 世纪 60 年代末和 70 年代初，麻省理工学院的杰罗姆·弗里德曼（Jerome Friedman）、亨利·肯德尔（Henry Kendall）和

来自于 SLAC 的理查德·泰勒（Richard Taylor）及他们的同事首次“瞥见了夸克”。这一瞥使弗里德曼、肯德尔和泰勒获得了 1990 年的诺贝尔物理学奖（图 4.53）。

图 4.53 理查德·泰勒（左），亨利·肯德尔（中）和杰罗姆·弗里德曼（右）手拉手庆祝获得 1990 年的诺贝尔物理学奖。在比约肯灵感的指导下，他们的电子同质子深度非弹性散射实验首次提供了存在夸克的清晰证据。图片来源：SLAC，美国物理学会艾米利奥·西格雷可视档案馆，《今日物理》收藏

像霍夫施塔德一样，弗里德曼、肯德尔和泰勒将电子射向静止的氢靶，然后在被散射的电子相对于入射电子的某个偏角上进行检测。但是他们所用的电子束要比霍夫施塔德所用的强 100 多倍，在这样高的能量下，大多数碰撞都是非弹性的，在过程中作为靶子的质子也被撞得粉碎。他们期望得到一些与霍夫施塔德类似的结果。但他们得到的结果却完全不同。

他们本以为会观察到随着探测光子的波长变短，散射截面也会

随之变小，就像霍夫施塔德在弹性散射的电子实验中看到的一样，这样就能够反映出质子的大小。但是他们吃惊地发现，散射截面的变化对波长变化的反应却相当迟钝，因此无法正确地记录质子的大小。这表示散射源并不依赖于质子的大小。

一位年轻的 SLAC 理论物理学家詹姆士·比约肯（James Bjorken，图 4.54）在这个方向上发现了有用的线索。比约肯花了两年的时间，试图理解在电子和质子散射过程中到底发生了什么。在那个时候，散射电子所看到的质子内部的未知组分，都是直接包含在结构函数当中的，本质上，结构函数描述了质子怎样区别于一个简单的点状粒子。总的说来，结构函数依赖于探测光子的波长，并且同入射电子经由光子传递给质子的能量有关。比约肯指出，能量和光子的波长互相纠缠，结合为一个复合量，该复合量既不具有质量的单位，也不具有能量的单位，实际上它仅仅是一个无量纲的数值。这看起来好像有点深奥，结构函数仅仅依赖于某个无量纲的数，这就表示不需要关注长度尺度。这种散射实际上是电子同一些带有部分质子动量的点状物体发生的弹性散射。

厨房里的水龙头可以帮助我们理解比约肯的想法。先慢慢地将水龙头打开，让水平滑地流下来。然后继续旋转水龙头让水流得更快，水流就会开始扰动。工程师们特别熟悉这种从平滑到扰动的跳变过程。这是一种变化，工程师们用一个叫做“雷诺数”的量来表示它，雷诺数是根据英国工程师兼物理学家奥斯伯恩·雷诺（Osbourne Reynolds）的名字命名的。雷诺数在大约 2000 以下时，水流是平稳的；当雷诺数达到 3000 以上，水流就开始扰动了。雷诺数为 2000～3000 时，两种情况都有可能发生。

准确地讲，雷诺数就是一个数字。它是流体的密度、流体的速度以及问题涉及的某些固有长度（可能是水管的直径）的乘积，再除以流体的黏滞度——即流体对运动的阻碍程度。证明以

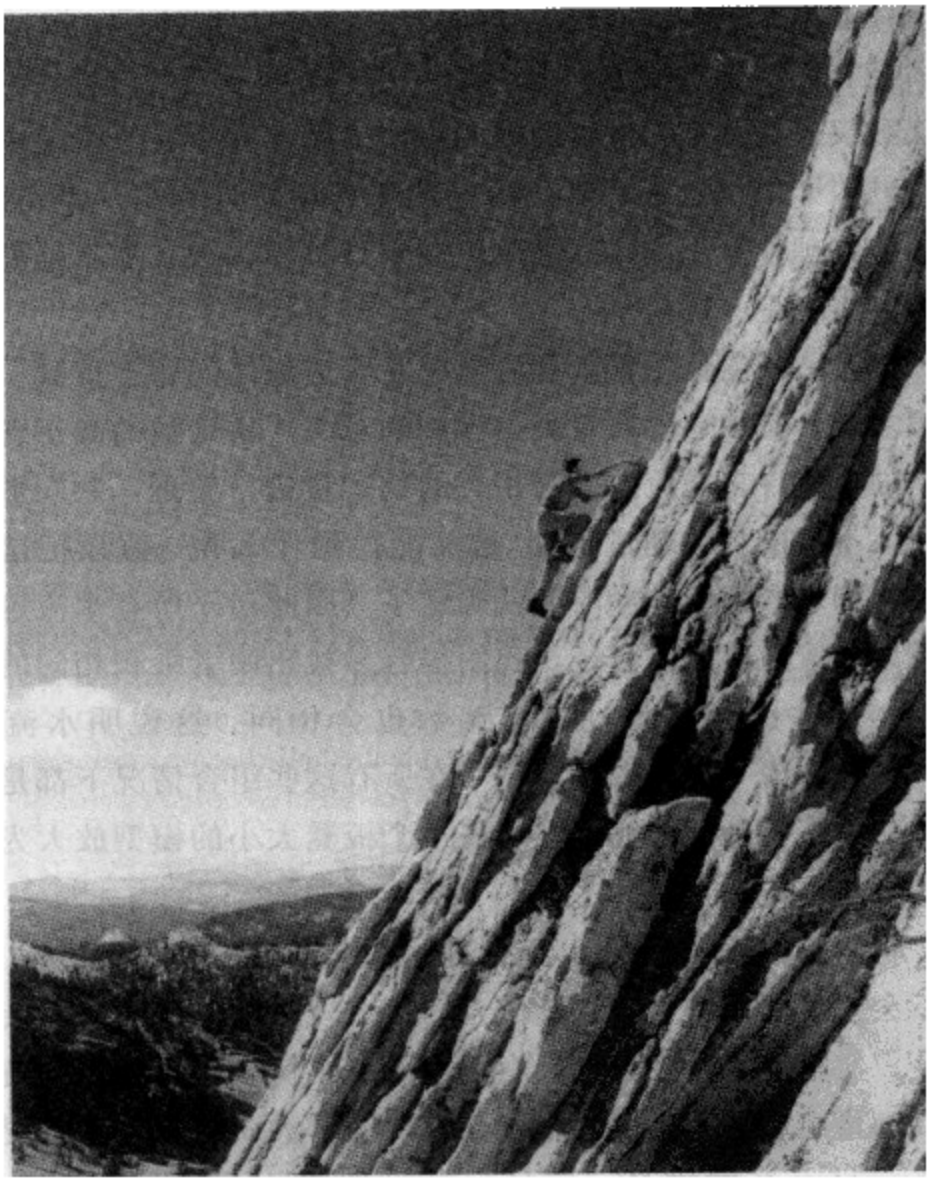


图 4.54　比约肯标度无关性：比约肯正在攀登约塞米蒂国家公园的大教堂峰，亨利·肯德尔拍摄于 1960 年。比约肯的革命性思想是揭示量子色动力学的里程碑，他的思想支持了以夸克为基础的电子-质子深度非弹性散射碰撞解释。图片来源：亨利·肯德尔，詹姆士·比约肯

上的结论只需要占用信封背面几行字的空白地方：稍微考虑一下

与流体流出水管有关的量，再确定好它们的量纲——米、秒等，然后把它们结合在一起。其背后并不存在什么令人着迷的理论。雷诺数仅仅是一个无量纲的数的事实表明，对于一个管道中的流体问题，如果一位工程师用古怪的工程单位计算出一个雷诺数，而一位物理学家用国际标准单位制来计算雷诺数，那么他们得到的数字将会是一样的。

但是除了能让工程师和科学家在管道流体问题上得到一致的答案之外，其中还隐藏着更为深刻的意义。雷诺数的真正含义在于，通过将许多与流体紧密相关的物理量进行整理，共同形成了一个无量纲的“物理量包”。这又能得到什么呢？假设把这种流体确定为水，密度和黏滞度就确定了。因此对于所有水流速度的范围以及管道可能的直径，我们很容易找出两者乘积相同的组合情况，这些组合所得到的雷诺数也会相同。这说明水流的性质——比如说平滑或者扰动——在所有这些组合情况下都是相同的。化学工程师们能够很容易将一个板凳大小的模型放大为真实大小的工厂。

将水管的直径放大 10 倍，并将水流的速度缩小 10 倍，雷诺数不会发生任何变化，它所代表的信息就是水流的情况不会发生改变。实际上，雷诺数是尺度无关的一个例子，尺度无关就是一个物理量组成的无量纲的组合，它说明扩大一个物理量就要相应地减小另一个物理量。这是一种在整个物理领域普遍有效的思想。

比约肯所做的就是，提出了一个描述电子和质子之间非弹性散射的尺度关系（标度无关性）。他所得到的尺度关系也是一个无量纲的数，可以同雷诺数作类比，它也像雷诺数一样不涉及基本的长度尺度。比约肯标度无关性（这是人们通常的说法）说明散射是通过点状的散射中心实现的。弗里德曼、肯德尔和泰勒发现了第一个存在比约肯标度无关性的证据，并且发现了核子内部

坚硬的部分。

理查德·费曼也开始用点状的质子组分来考虑问题，他用了一个简单得令人吃惊的模型，但直到现在，粒子物理学家们仍然很好地用着这个模型。他在核子中引入了部分子（partron）的概念——它们是一些能自由运动的点状散射中心，并以此来理解高能强子之间的碰撞。一个具有很大动量的核子（或者其他种类的强子）中，部分子通过某种方式在它们之间分配动量，这是核子的一种性质，而不是人为的假设。这些部分子所携带的核子动量分数一般用 x 表示，因此对于任何一个部分子，x 的取值范围是 0～1。在深度非弹性散射的情况下，动量分数 x 实际上就是比约肯得到的无量纲数值的倒数。

关于尺度问题以及点状物的散射问题，费曼和比约肯得到了相同的结论，但是费曼的部分子模型还额外提供了一种简单而直观的图像，用它可以描述电子撞击质子的情形：入射电子发出的光子同点状部分子相撞（图 4.55）。在 SLAC 观测到的非弹性散

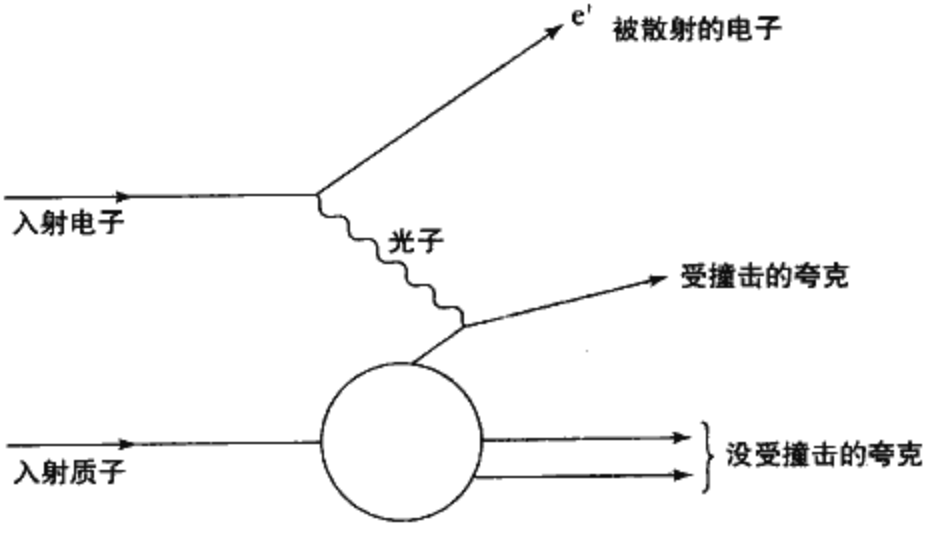


图 4.55 部分子模型中的电子-质子深度非弹性散射。受击打的夸克所携带的动量占入射质子的比例为 x

射仅仅是所有部分子发生弹性散射贡献的总和。在强子-强子碰撞过程中，例如质子-质子碰撞过程，一个质子中的部分子会和另一个质子中的部分子发生弹性散射，图 4.56 用胶子清楚地说明了该过程。

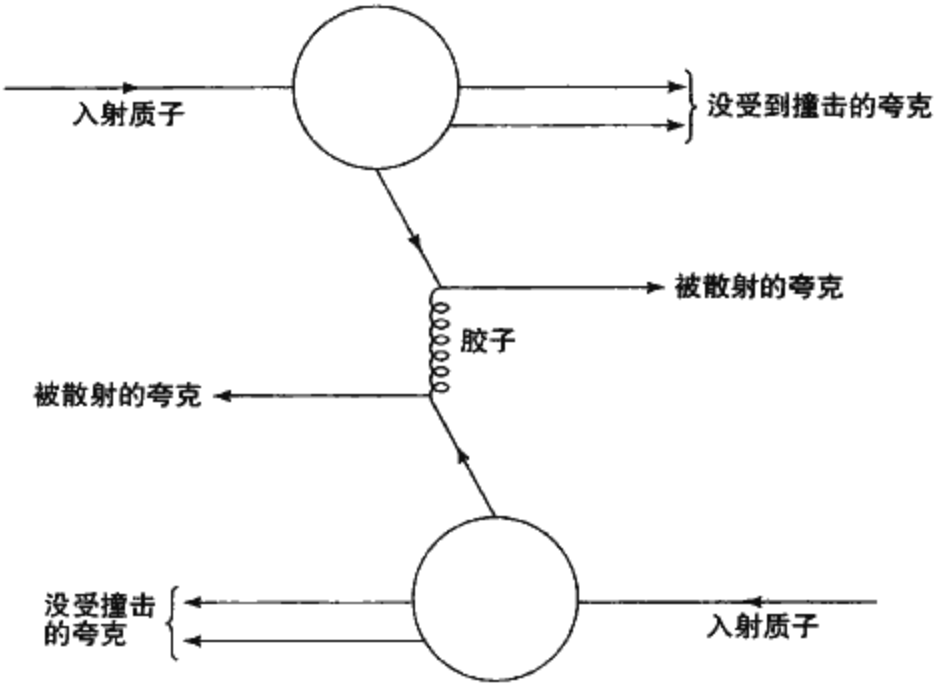


图 4.56　两个质子碰撞的方式之一，图中用部分子模型形象地加以表示

要完整描述电子同质子以及中子发生的非弹性散射过程，不只需要一个结构函数，而是需要两个，才能足以容纳所有光子的极化情况。比约肯的工作表明，这两个结构函数都依赖于动量分数 x，卡兰和格罗斯在这一基础上经过努力，通过假设这种坚硬的散射中心具有 1/2 的自旋，将两者联系起来。早在 1969 年，理查德·泰勒就在英国利物浦的一次会议上发表了支持卡兰-格罗斯关系的数据。随后出现的更多 SLAC 数据显示，电子同质子散射和电子同中子散射有所不同。CERN 的实验者们也适时地进行了中微子同核子的散射实验，这是一种由 W 玻色子代替光子传

递相互作用的深度非弹性散射过程，见图 4.57。从今天的观点来看，通过综合各种数据得出的结论看起来很明显：这些点状的散射中心就是夸克。

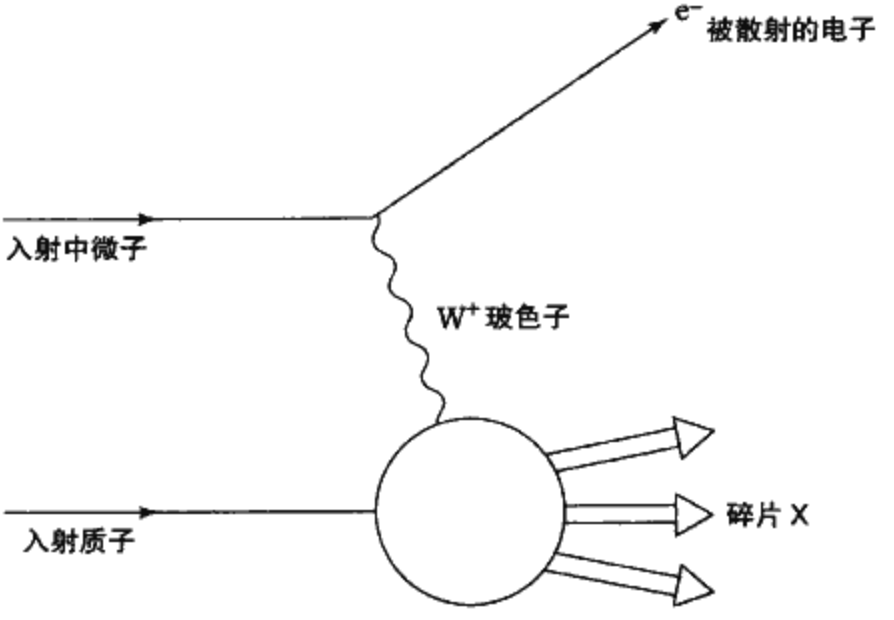


图 4.57 中微子-质子深度非弹性散射。入射（电子）中微子放出 W^+ 玻色子将质子击碎

如此说来，结构函数就是核子中夸克所具有动量的贡献。这体现在：至今，几乎所有的粒子物理学家都了解质子和中子的结构。测量结构函数，并且尝试用 QCD 理论来解释这些测量结果，是一项巨大的工程。HERA 之所以如此重要，就是因为它能实现电子和质子对撞，实验者们能够看到夸克携带有一部分核子动量，而这是在其他任何加速器上看不到的。用行话讲，它们能够探索低 x 区域。

CERN 的中微子散射结果于 1973 年发表，其中中微子通过 CERN 的质子同步加速器产生，并通过伽格米勒气泡室（Gargamelle）进行探测，1973 年正好处在理论物理学家们努力发展 QCD 理论的时期。然而即使在那个时候，伽格米勒气泡室的小组

还是马上就确认，有大约50%的核子动量由胶子携带，该实验结果后来又被其他中微子实验小组证实。人们从一开始就明白，三种所谓的价夸克——即那些对应于与粒子结构有关的夸克模型的夸克——并不完整，因为从它们并不能得出结构函数的测量形式。这三种价夸克通过胶子相互作用，现在已经在实验中观测到了，这些胶子反过来又能够生成夸克-反夸克对。根据这种图像（这也是现在的基本观点），核子中的夸克处在翻腾的胶子和夸克-反夸克海中。实际上，伽格米勒气泡室那些早期的结果已经显示了核子中存在反夸克的证据。

通往夸克自由之路

当实验物理学家们还在搜集标度无关性的证据时，费曼和比约肯已经开始支持夸克-部分子的观点，特霍夫特和维尔特曼也在证明，整个非阿贝尔规范理论的神奇家族确实是重正化的。格拉肖、温伯格和萨拉姆对弱力的解释现在可以被当作基础理论，杨-米尔斯理论忽然就变成了一个寻找强力理论的好地方。标度无关性成为了解杨-米尔斯理论在高能状态行为的催化剂。

标度无关性确实推动了格罗斯和威尔茨克的研究。威尔逊当时进行了有关重正化群的工作，卡兰和西曼吉克对其进行了改造，为了探究杨-米尔斯理论在高能状态下的行为，格罗斯、威尔茨克和波利茨在他们的工作中找到了一套理想的工具——重正化群的时代降临了。关于非阿贝尔理论，格罗斯和威尔茨克在他们1973年那篇有关渐近自由的文章中写道："我们已经发现它们具有逐渐接近自由场的重要特性，这或许是可重正化的理论特有的。"这就是渐近自由的含义，在高能情况下，组分之间的相互作用趋近于零，它们之间没有相互作用；换句话说，就是像"自由场"。他们继续写道："这种渐近自由理论将体现出……比约肯标度无关性，到目前为止，场理论还解释不了比约肯标度无关

性，因此我们建议，应该考虑一种能解释它的有关强力的非阿贝尔规范理论。”全世界都在倾听、并且实施他们的建议。

标度无关性是在SLAC的深度非弹性电子-质子散射中首先被发现的，这是一种令人非常吃惊的实验结果，它对渐近自由的出现以及核子中的点状“小块”参与散射的思想都十分重要。人们最终证明这些坚硬的小块就是夸克，从此夸克变成了一种更真实的物体，而不仅仅是解释强子质量分布的一种工具。渐近自由是在颜色的基础上描述夸克之间相互作用理论的关键性质，它说明在适当的条件下，夸克的行为就像自由粒子。关于这一点存在许多相关的线索和假设，但却没有证据，因为尽管电子能看到那些夸克聚集在核子中，但夸克永远不可能被单独看到。但在封闭夸克的强子牢笼之外，会存在自由夸克吗？它们是不是正等着那些采用正确方法去寻找它们的人呢？粒子物理学必须知道。

寻找自由夸克

哲学家们大概会对“寻找一些你认为不可能找到的东西”这个逻辑问题有很多话要讲。毕竟，如果你希望找到某个东西，并且确实找到了，那么你就可以举行一场酒会来进行庆祝。但是如果你认为你根本不能找到某个东西，并且你在实验中确实也没找到它，在你的酒会上可能就会有许多人在想，也许你没有用合适的实验去寻找它，或许你期望找不到的那个东西确实存在于某个地方，只是它藏在一个你没找到的地方罢了。

夸克的情况也是这样。现在没有人会真的期望能找到像实验物理学家们所熟悉的自由电子或π子那样的自由、孤立的夸克。这使得寻找自由夸克的实验做起来相当乏味，这不仅是因为存在逻辑上的问题——你不能仅因为实验上没有找到自由夸克，就推断说它们确实不存在，而且一想到你期待从实验仪器中得到一个

什么也找不到的结果，就会更加令人扫兴。这真是让人有点沮丧。

然而人们或许也能从中得到补偿，随着时间的推移，夸克禁闭的思想在人们心目中已经越来越根深蒂固，因此如果能找到自由夸克，意义将更加重大。无论谁找到了自由夸克，他都应该马上开始准备他在诺贝尔奖获奖仪式上的演讲稿了，而且他的发现将给理论界带来很大的冲击。因此这是一场豪赌。

人们做了大量各式各样的实验来寻找自由夸克。从夸克概念诞生那一刻起，搜寻自由夸克的旅程就开始了。第一次实验是在1964年，CERN质子同步加速器和布鲁克海文国家实验室的AGS质子加速器都拍摄了气泡室照片，在那时它们都是能量很高的加速器。毕竟盖尔曼自己提出了用“能量最高”的加速器来证明“真实夸克”不存在的想法。并且，可以肯定的是，第一批探索者们一无所获。

但盖尔曼也提到一种想法：宇宙射线可能会使地球表面形成一层稳定的夸克污染物。这至少在整个实验领域内激励了其他一些实验的出现，这些实验立刻被冠以非加速器实验，或者粒子物理反传统实验之类的名字。

据盖尔曼回忆，密歇根大学的光学研究员彼得·弗兰肯（Peter Franken）常常会在午夜给他打电话，向他报告在牡蛎中寻找夸克的进展，弗兰肯相信自由夸克会集中出现在海水中。但最著名的在大量物质中寻找自由夸克的实验，可能还是由斯坦福大学的威廉姆·费尔班克（William Fairbank）和他的同事做的实验，他在1977年报告说他发现了带分数电荷的粒子，而且很有可能是夸克。

费尔班克实验是美国物理学家罗伯特·密立根（Robert Millikan）经典工作的翻版，密立根著名的油滴实验使他获得了1923年的诺贝尔奖。密立根成功测量了基本电荷单位，他用电场

限制带电的小油滴使其静止，电场对油滴施加向上的拉力，与向下的引力达到平衡。小油滴在喷雾器中诞生的时候会得到一个静电荷，密立根和他在芝加哥大学的同事经过几年的时间，测量了成千上万的油滴，发现油滴所带的电荷总是一个基本电荷单位的整数倍。他们的实验结果既给出了电子电荷的值，也说明了电荷是量子化的。费尔班克和他的同事用小超导铌球代替了油滴，并且用磁悬浮来平衡地球引力。在施加震荡电场并精确监测了每一个小球的运动情况之后，他们宣称探测到+1/3和-1/3倍的分数电荷。但是所有想再次得到他们的实验结果的尝试都失败了。与费尔班克实验最为神似的工作，是热那亚大学的吉亚卡摩·莫波格（Giacomo Morpurgo）和他的同事做的悬浮钢球和石墨颗粒的实验，他们的实验从1966年开始，坚持了15年，但没有观测到分数电荷的一丝迹象。最普遍的观点认为，费尔班克的实验是外部影响的牺牲品，其结果可能是由一些磁场背景所导致的现象，他可能根本就没有看到自由夸克。

对于费尔班克和他的小组来讲，找到分数电荷是他们实验存在的理由。因此回顾一下密立根于1910年发表在《哲学杂志》（*Philosophical Magazine*）上的一篇报告中所做的评论是很有意思的事情。密立根写道："我已经放弃了有关某滴油滴呈现出的一个不确定并无法重复的观测结果，该带电油滴所给出的电荷值明显比e（电子电荷）低约30%。"如果不管正负号，一个比电子电荷低30%的电荷正好是一个上夸克所带的电荷——这是不是一次侥幸的成功，是不是密立根真的发现了什么？不太可能，密立根意识到那个出问题的液滴很可能是一个水滴，它可能由实验中的蒸汽产生。实际上他推断，有问题的液滴很可能会快速蒸发掉，因为它们都特别小。第二年，他在《物理评论》上发表了另一个实验结果，新实验使用了油滴（以及其他一些防水蒸气材料构成的液滴）和改进后的仪器。他在每一次测量中，发现液滴所

带的电荷都恰好是某个最小电荷量的整数倍。因此密立根的工作确实没有给自由夸克的存在提供足够的证据，但它确实体现出密立根是一位最优秀的实验物理学家。

密立根油滴实验的精神流传了下来。1996 年，由马丁·佩尔（Martin Perl）带领的一个 SLAC 小组，再次公布了利用油滴实验寻找分数倍电荷的实验，这是三种夸克之间差别最大的性质，也是所有寻找自由夸克的人赖以识别夸克的标志。在这次 SLAC 实验中，油滴先掉进上方水平板上的一个小洞，在进入放在下面的一个与之平行的水平板上的小洞之前，还要经过一层 1 厘米厚的空气，见图 4.58。每个液滴的直径大约都是 7 微米，由一台定做的滴液器产生，其生产速率为每 2 秒钟产生一滴液滴。两个板之间有一个震荡电场，它可以根据液滴所带的电荷改变每个液滴的速度和电场的上下方向。该小组利用频闪观测仪和摄像机自动监测了大约 600 万个液滴，得到了每个液滴的直径、在空气中的末速度以及它的电荷。实验者们一共检测了 1.07 毫克的油。实验显示，如果存在电荷为 $\pm 1/3$ 或 $\pm 2/3$ 的孤立夸克，那么在 95%的可信度下，存在孤立夸克的数量比存在核子数量的 2.14×10^{-20} 还要少。

这个结果和不存在自由夸克是不一样的。它仅仅表示如果存在自由夸克，它们的量会特别的少。该结果为夸克的存在设置了一个上限，这是实验物理学家们表示他们寻找某些东西但没找到时所用的典型表示方式。这当然也是表达搜寻夸克结果的普通方式，可信度表示统计的不确定性，在粒子物理学中，对于大多数实验 95%的可信度一般都是可以接受的水平。可以将自由夸克与核子数的比例形象地表示出来，大致相当于有 95%的把握说明，这个比例比一粒“调皮的”小沙子同一堆体积超过 200 万平方米的胡夫金字塔那么大的沙堆的比例还要小。

还有许多在大量物质（例如在铁、水银以及石墨）中寻找自

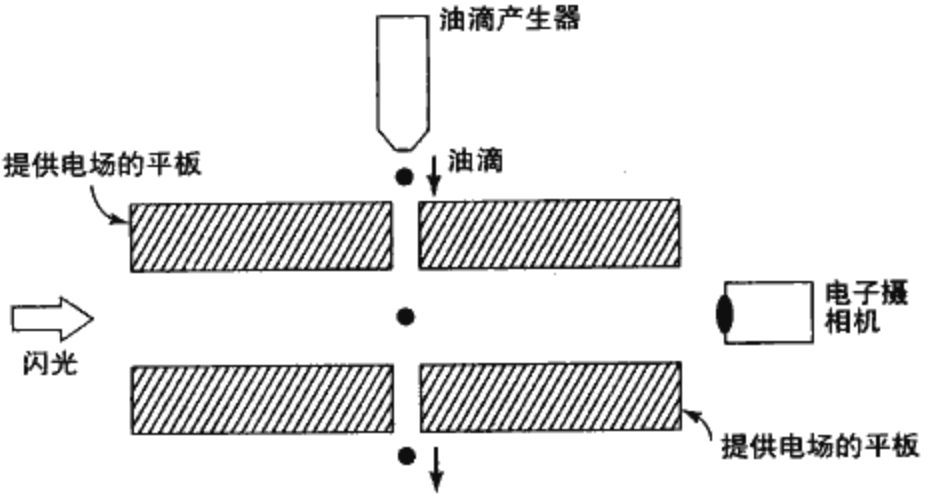


图 4.58 SLAC 小组重新进行经典的油滴实验，试图寻找油滴上的分数电荷

由夸克的其他实验，有一些小组还尝试了铌。匹兹堡大学的尤尼尔·皮里拉（Unil Perera）和他的同事通过在硅中进行寻找，并且假设夸克表现为一种带分数电荷的掺杂物，最终给出了与 SLAC 的油滴实验相类似的限制。在该实验中，半导体硅所带的不同电荷都处于束缚态，但束缚得很弱，任何自由夸克都可以通过用光照射材料来发现。可供选择的另一种大量物质是海水，但并不是每个人都能像弗兰肯那样想到用牡蛎来浓缩夸克。一个由来自于牛津附近卢瑟福-阿普尔顿实验室（RAL，Rutherford Appleton Laboratory）的英国物理学家，以及来自伦敦的南安普顿玛丽王后大学（the University of Southampton and Queen Mary）和韦斯特菲尔德学院（Westfield College）的化学家所组成的合作小组，在 1992 发表了在海水中的搜寻结果。他们用蒸发法在铌球外包裹了一层海水残余物——主要是盐，他们的想法是，由于宇宙射线的作用，自由夸克可能会沉积在海水里，通过蒸发可能会增大任何带有分数电荷的组分浓度，然后再用带电板之间的磁悬浮来测量裹着外衣的小球是否带有分数电荷。实验者们还是

没有找到自由夸克，并确定了每克海水小于 0.001 个夸克的比例上限。

一些人还检测了月球岩石，但也无功而返。RAL 海水小组的三个成员彼得·史密斯（Peter Smith）、格雷厄姆·霍默（Graham Homer）和沃利·沃尔福德（Wally Walford）也是更早的一个寻找自由夸克的合作小组的成员，该小组试图在陨石（这些原始而未被破坏的材料来自于我们太阳系的诞生时期）中寻找自由夸克，但是他们什么也没发现。在 2002 年，佩尔和他的小组发表了一种由他们的油滴实验改造出的实验，他们将一些有着 500 万年龄的陨石放入油滴中，但同样没有任何结果。

一个寻找夸克的著名场所是宇宙射线，一些宇宙射线具有的能量比任何人造加速器加速的粒子都要高得多，因此就存在看到在加速器中看不到的东西的可能。已经有 50 多个在宇宙射线中寻找自由夸克的实验结果公开发表，其中包括著名的自由夸克“候选人”，它是由悉尼大学的布莱恩·麦克卡斯科（Brian McCusker）于 1969 年在云室照片中发现的，但它只是经不起时间考验的一次性事件。20 年过去了，一个由美国和日本科学家组成的小组为了寻找分数电荷，在东京以西 300 千米处的神冈矿井底部搜集了 3 年数据。他们使用通常被用做观测质子衰变的 Kamiokande Ⅱ 探测器。该探测器含有 2400 吨水，有将近 1000 个光电倍增管用于观测，光电倍增管是将微弱光信号转化为电脉冲的典型探测仪。该探测器同其他几部质子衰变探测器一样，被安放在矿井底部，因为周围的岩石能够成为阻挡额外宇宙辐射的良好屏障。因此当光电倍增管发现了什么东西，就很可能是由大量水中存在的质子发生的质子衰变事件，也有可能是经过了一个自由夸克。尽管实验者们积累了 7000 万次事件，但大多数都是速度非常快的 μ 子，他们没有找到自由夸克。

当然，大型的加速器也投入了搜寻夸克的工作。通过在 LEP

上观察电子-质子碰撞，ALEPH 探测器合作小组分析了 18 万次 Z 粒子衰变为强子和轻子的过程，从而寻找与众不同的质量或电荷信号。两年之后，在 1995 年，OPAL 小组公布了另一个搜寻结果，这一次他们观测了超过 100 万次 Z 粒子衰变为发生强相互作用粒子的过程。两个小组都没有发现自由夸克，但两个小组都把产生自由夸克的上限范围又缩小了。其他一些实验者利用 Tevatron 在能量高达 1.8 太电子伏的质子-反质子对撞中寻找自由夸克，但也什么都没有发现。

其实，如果任何一种东西具有分数倍的电荷，那都会是一条特大新闻。毕竟也有可能存在带有比电子所带电荷单位还要小、但不带颜色的粒子，但它们不是夸克。因此最好将搜寻自由夸克的工作称为“带分数电荷粒子”的搜寻。不管它是什么粒子，至今还没有人发现任何能够经得起第二次观测的东西。但实验物理学家们必须也确实应该继续寻找下去。

更多的夸克

科学家们不需要走很远的路去寻找 u 夸克和 d 夸克，因为他们就是由这两种夸克组成的。构成世界的物质原料是原子，原子的核是由中子和质子组成的，而质子和中子又是由 u 夸克和 d 夸克组成的，只是我们很难见到它们。

因此，这两种夸克就坐在实验室的凳子上等着人们找到它们。第三种夸克——s 夸克的发现，是超自然的介入同传奇性的科学结合的产物，即该结果来自于宇宙射线和山峰顶端的实验。奇异粒子——明确地说是奇异夸克——令每一个人都感到吃惊。但等到第四种夸克被发现时，至少有一些物理学家已经做好准备等着这一刻的降临。

比约肯和格拉肖那时一起在哥本哈根工作，他们在 1964 年

的一篇论文中引入了一个被他们称为粲量子数（charm）的新量子数，该论文意在通过加入另外一种基本“砖块”来扩展八正法。6 年之后，格拉肖在哈佛大学和约翰·里尔普洛斯、卢西亚诺·迈阿尼（Luciano Maiani）一同指出，第四种夸克——粲夸克将会优美地解决弱相互作用中一种令人气恼的问题。例如像 $K^0 \to \mu^+ + \mu^-$ 以及 $K^+ \to \pi^+ + e^+ + e^-$ 这样的衰变过程看起来是允许的，但实验物理学家们却确定它们一般不会发生。不仅是这些，存在第四种夸克还表示夸克和轻子满足格拉肖和他的同事称为“显著对称性”（remarkable symmetry）的性质：电子和电子中微子可以通过上夸克和下夸克和谐地联系在一起，μ 子和 μ 子中微子则依赖于奇异夸克和粲夸克。实际情况是，如果不将粒子联系到一个家庭中（包括带颜色的夸克），标准模型就不成立——为了保证可重正性，就需要正确的粒子组合方式。

4 年之后，粲夸克在 1974 年被发现了，有两个小组同时公布了他们的结果。利用位于 SLAC 的新 SPEAR，伯顿·里克特（Burton Richter）和他的同事见证了在电子-质子碰撞产生强子的过程中，散射截面存在一个陡峰，见图 4.59。这个峰特别陡，因此很容易被漏掉，在逐步增加加速器的能量时，很可能将该峰所处的能量夹在两个设定的能量之间而忽略。他们将这个峰的出现解释为一种共振态，称之为 ψ 粒子。而在大洋彼岸的美国，丁肇中（Sam Ting）和他的小组在布鲁克海文国家实验室的 AGS 也发现了一个共振态。他们的实验是通过将一束质子打到铍靶上，从而观测在质子-质子碰撞过程中所发出的电子-正电子对的总质量。从一个质子产生的一个夸克与从另一个质子产生的一个反夸克发生湮灭，生成光子，最终生成一个电子-正电子对，或者某些其他的轻子对。该过程是业内人士都很熟悉的 Drell-Yan 过程，见图 4.60。丁肇中的小组发现的共振态和里克特所发现的一样，也位于相同的能量，只是他们把它叫做 J 共振态。这两个小组在

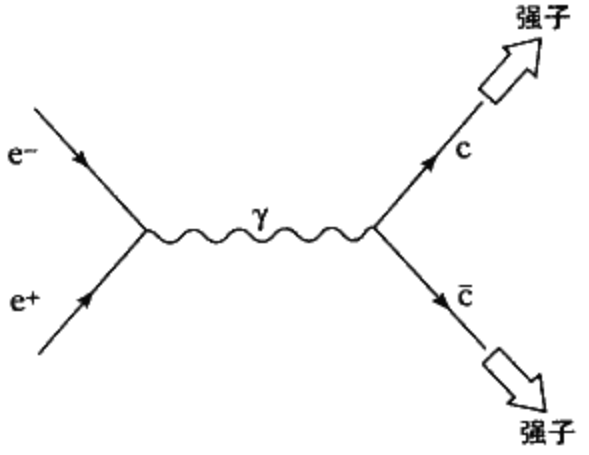


图 4.59 粲夸克的发现：在 e^+e^- 湮灭过程中，位于 3.1 吉电子伏能量处的一个共振态表明诞生了粲介子 ψ，它是一种 $c\bar{c}$ 对

该共振态的名字问题上始终无法达成一致，从此粒子物理不得不将其称为 J/ψ 粒子。J/ψ 粒子是最简单的一种介子，包括一个粲夸克和一个反粲夸克。

丁肇中和里克特由于在这次发现中的贡献，获得了 1976 年的诺贝尔奖，这次发现令粒子物理界大为震惊。1974 年的那个 11 月有时也被称作“十一月革命”，因为它永远改变了粒子物理学之路。新共振态的能量范围窄得不可思议，最终人们只能将其解释为存在第四种新的夸克，J/ψ 粒子很难在短时间内衰变为 π 子，用胶子概念和 QCD 理论能够自然解释这种现象。这是 QCD 理论早期能得到人们认可的证据之一。实验物理学家们有一套全新的理论可以利用了，他们可以将新发现的粲夸克和已知的夸克组成粒子，或者以简单的粲夸克-反夸克对为基础组成其他的态，它还帮助人们坚定了对年轻的 QCD 理论的信心。在发现 J/ψ 粒子仅 10 天之后，第一个可能的粲夸克-反粲夸克对就在 SLAC 被观测到了，它是一种由粲夸克和反粲夸克组成的“激发态”，现在人们知道，它就是质量为 3686.0 兆电子伏的 ψ（2S）粒子。第

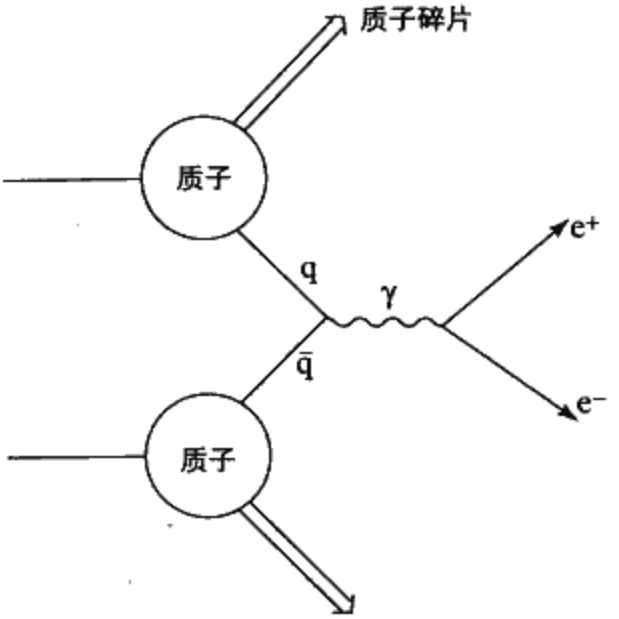


图 4.60 粲夸克的发现：来自于一个质子的夸克同来自于另一个质子的反夸克发生湮灭。得到的产物被称作 J 共振态

一个粲夸克重子的迹象来自于 1975 年布鲁克海文国家实验室的中微子-质子散射实验。就在 2002 年，费米实验室的 SELEX 实验公布了一种可能包含 2 个粲夸克的重子——这在 4 种味道的夸克模型中是十分重要的，见图 4.61 和图 4.62。但是最早出现“纯粲荷”的明确证据——一个没有反粲夸克伴随的粲夸克，是 1976 年在 SLAC 第一次发现 D 介子时出现的。随着 D^0 介子（一种由粲夸克和反上夸克组成的介子，或者表示为 $c\bar{u}$）被观测到，同一年内 D^+ 和 D^- 介子也被观测到，粲夸克存在的事实已经毋庸置疑。

20 世纪 70 年代是粒子物理的一个黄金时期。就像为了要让结论更加确定，人们在 1976 年还得到了另一个戏剧性的发现。佩尔和他的同事分析了大约 35000 次 SPEAR 上的电子-正电子湮灭事件，仅仅有 24 次事件无法用当时已知的粒子明确地解释。这些事件的形式是 $e^+ + e^- \rightarrow \mu^- + e^+ +$ 损失的能量，或者是

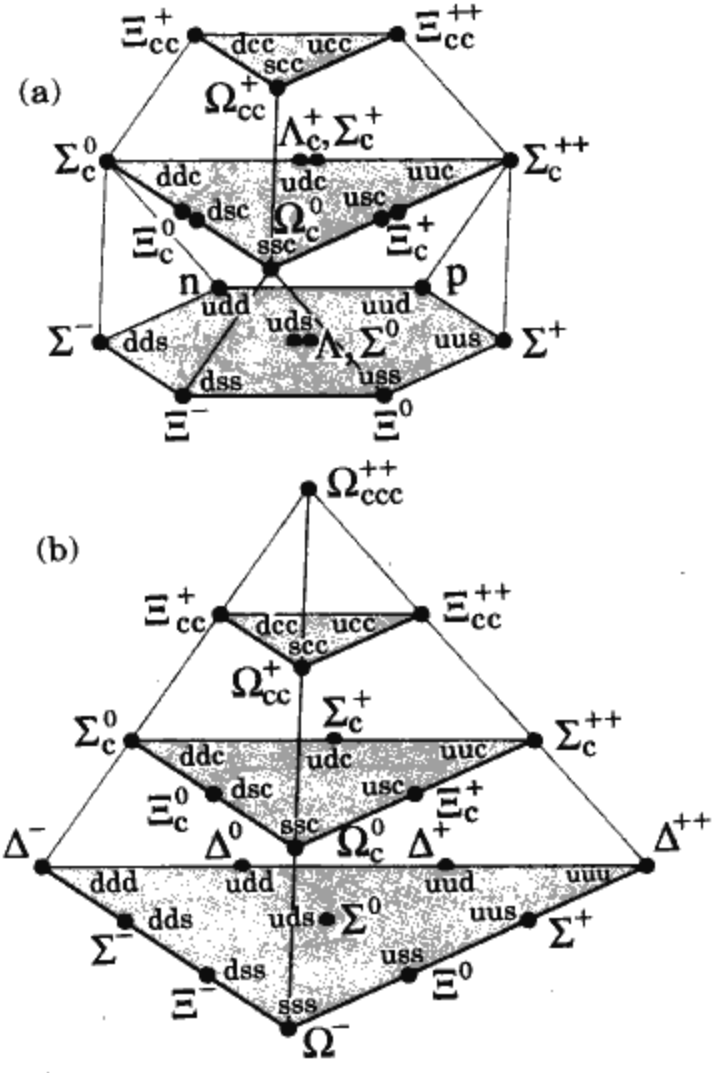


图 4.61　新增的粲夸克扩展了重子家族；这些庞大家族的“底层”是大家比较熟悉的那些粒子，它们仅由 u 夸克、d 夸克和 s 夸克构成。图片来源：粒子数据小组

$e^+ + e^- \to \mu^+ + e^- +$ 损失的能量，这同人们预料的情况一样——电子-正电子的偶湮灭会产生一对较重的轻子。佩尔和他的同事观测到了一种全新的轻子，即陶轻子，或者叫 τ 轻子。他们发现的过程是 $e^+ + e^- \to \tau^+ + \tau^-$，接下来 τ 子会衰变成数量大致相同的电子或 μ 子，负 τ 子相应的衰变过程分别是 $\tau^- \to e^- + \bar{\nu}_e + \nu_\tau$

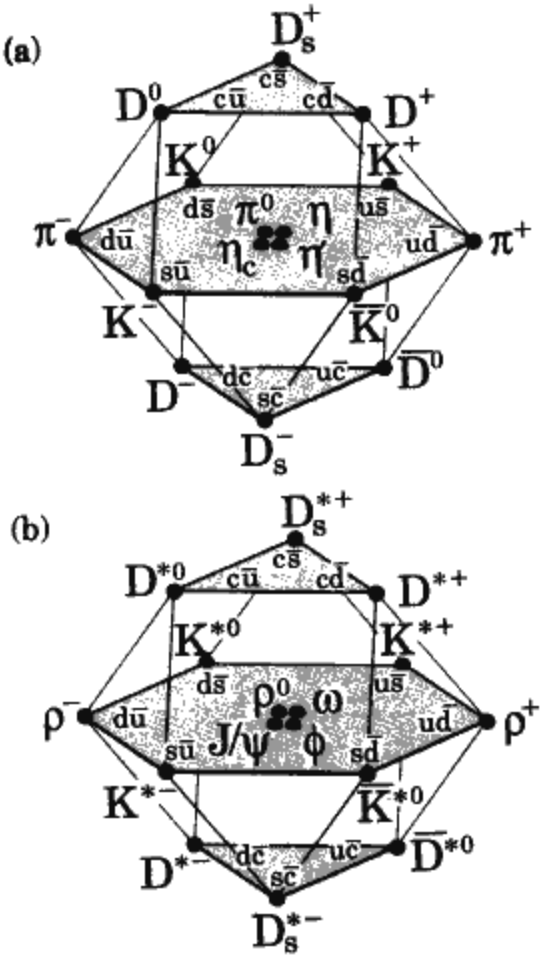


图 4.62　粲夸克扩展了介子家族。中间的一层是人们比较熟悉的介子家族。图片来源：粒子数据小组

或者 $\tau^- \to \mu^- + \bar{\nu}_\mu + \nu_\tau$。

τ 子与电子和 μ 子比起来非常重。它的质量是 1777.1 兆电子伏，这个数大约是 μ 子质量 105.7 兆电子伏的 17 倍，与“轻如鸿毛”的电子质量 0.511 兆电子伏相比，更是大得离谱。τ 子像电子和 μ 子一样，也是一种“真正”的基本粒子，也有与之相伴的中微子，即 τ 中微子，或称 ν_τ。夸克与轻子之间的对称性还提供了一种暗示，正如电子是一个包括上夸克和下夸克的“家庭”的一部分，μ 子是一个包括奇异夸克和新发现的粲夸克的“家庭”

的一部分一样，τ 子也应该同另外两种夸克组成一个家庭。换句话说，τ 子的存在表示还有另一对夸克存在，它预示了另一个夸克一轻子家庭的诞生。

没过多久，人们就迎来了下一种夸克——那两种夸克中的一种。没有人知道这种新夸克的质量会是多少。丁肇中用来测定粲夸克的方法，看起来是着手进行研究的一种好办法，里昂·莱德曼（Leon Lederman）和他的同事，利用费米实验室那台 400 吉电子伏的质子同步加速器进行了实验，该实验在本质上是丁肇中实验的翻版，但实验的能量更高。在 1977 年，他们在大约 9.5 吉电子伏的地方发现了一个共振态，他们将其称为“宇普西龙子”（upsilon），即 γ 子。这已经超出了 SLAC 的 SPEAR 的能量范围，如果不是这样，SPEAR 早就会在电子-正电子碰撞实验中夺得发现 γ 子的荣誉。1977 年 6 月 30 日，在费米实验室研讨会上公布发现 γ 子的那一天，PLUTO 合作小组当时正在用他们的探测器监测 DESY 的 DORIS 存储环中的碰撞，他们提交了一项升级 DORIS 能量的提议。幸运的是，随着发现 γ 子的结果公布，他们迅速地将他们的冲击能量提高到了能包括 γ 子的范围。DESY 的实验者们在第二年就看到了 γ 子。

美丽的底夸克

γ 子的质量为 9460.4 兆电子伏，大约是 J/ψ 粒子质量的 3 倍，它也由夸克和反夸克对组成，只不过这次涉及的夸克是 b 夸克或者叫底夸克（bottom quark），有时也称之为美夸克（beauty quark）。在 γ 子被发现后不久，人们就首次观测到了由同一组夸克-反夸克对组成的能量更高的态，最先发现的是一种表示为 γ（2S）的 $b\bar{b}$ 激发态，这同它们较重的表亲 J/ψ 粒子的情况极其相似，见图 4.63。这种令人欣喜的性质被称作“裸底荷”，其存在的证据是在 1981 年由 CERN 的 ISR（Intersecting Storage Ring

proton-proton collider，交叉存储环质子-质子对撞机）和 CLEO 合作小组利用 CESR（the Cornell electron-positron collider，康奈尔电子-正电子对撞机）首次发现的。CERN 的小组观测到了“底重子” Λ_b 的迹象，它是一种由上夸克、下夸克和底夸克所组成的复合粒子，其质量为 5624 兆电子伏，至今仍是最重的一种重子，也是唯一一种得到证实的底夸克重子。它极短的寿命是一个谜。几乎在他们发现 Λ_b 粒子的同时，CLEO 合作小组首次发现了 B 介子的迹象，B 介子中只有一个夸克是底夸克。最先发现的一批 B 介子是 B^+、B^- 和 B^0 粒子，它们的夸克组分分别是 $u\bar{b}$、$\bar{u}b$ 和 $d\bar{b}$。这 3 种粒子的质量都是 5279 兆电子伏，是质子质量的

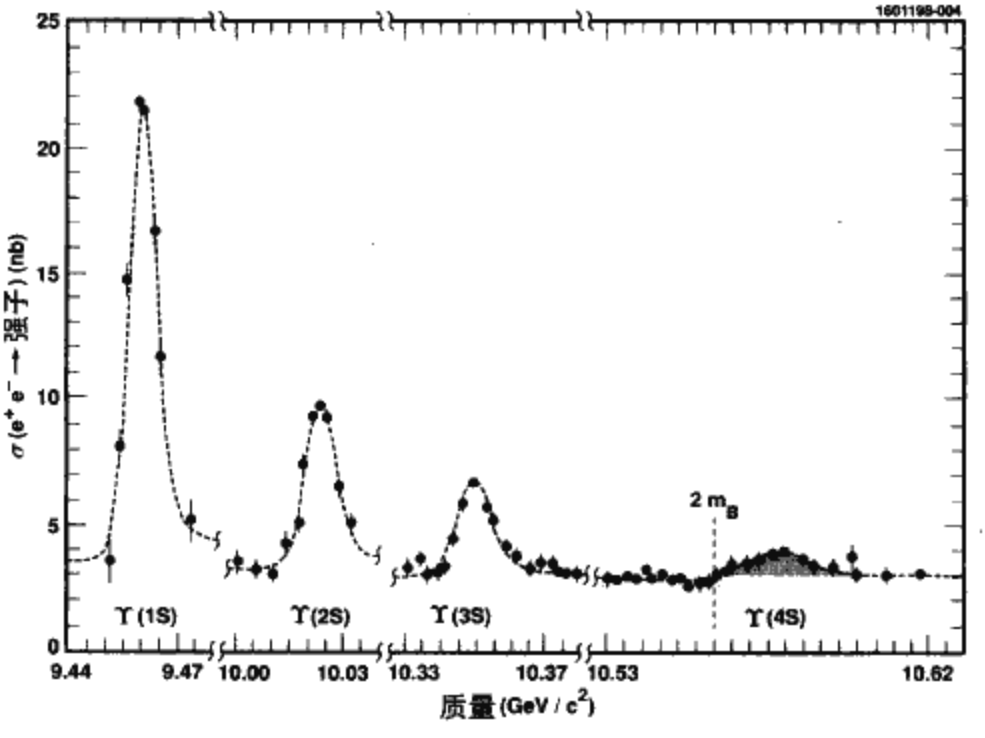


图 4.63　γ 粒子。γ 子是底夸克和反底夸克的一种束缚态，它有许多激发态，可以通过调高加速器的能量来产生，这些激发态在散射截面点图上看起来像是一些小突起。图片来源：康奈尔大学，哥伦比亚大学——石溪（CUSB）合作小组

5倍多，底介子们确实是一个沉重的粒子种类。

实验物理学家们在发现基本的D介子、B介子以及J/ψ粒子和ϒ粒子之后，又发现了一整套“重夸克”介子，换句话说，就是那些包含c夸克和b夸克的介子。有像B_s^0（夸克结构为$s\bar{b}$）这样由奇异夸克和底夸克组成的介子，有像D_s^+（$c\bar{s}$）和D_s^-（$\bar{c}s$）这样由粲夸克和奇异夸克所组成的介子，还有一系列由各种基本夸克组成的激发态粒子。此外，实验者们还发现了许多粲重子和底重子。

1998年5月5日，在一次费米实验室的研讨会上，CDF（Collider Detector Facility at the Teratron，Teratron对撞机探测仪）小组的科学家们公布，他们已经发现了B_c介子，它是一种由$c\bar{b}$组成的粒子，这是人们第一次观测到由粲夸克和底夸克所组成的粒子。它的质量大约是6400兆电子伏，其寿命大约是0.64皮秒，但最引以为豪的是，它是最后一种“普通”介子（其原因会在适当的时候加以说明）见图4.64。“普通”是从介子是由夸克-反夸克对组成的意义上来讲的，其他介子都是由这种组合组成的。实验者们花了50年，终于找到了所有的普通介子。

由于由粲夸克和底夸克组成的粒子都特别重，可想而知，它们衰变为较轻粒子的可能性也非常大，并且衰变也十分复杂。例如，每一种基本的粲介子（像D^0介子、D^+介子和D^-介子）以及每一种基本的底介子（像B^0介子、B^+介子和B^-介子）都会有几十种不同的衰变途径，一般是衰变为各种各样的K介子（kaon）和π子。仅B介子家族自己，总共就有大约1000种衰变模式。许多这类信息都是用衰变率来表示的，衰变率就是衰变可能性，它是衰变粒子的“散射截面”。衰变率通常称作宽度，用Γ表示，它是粒子寿命的倒数。产生某种特定产物的衰变用相对于总衰变可能性（也就是衰变总宽度）的形式表示。例如B介子的衰变过程$B \to \bar{D}^0 + l^+ + \nu_l$，其中$l^+$表示任何一种带正电的轻子，

$u\bar{u}$ π^0,η,η'	$u\bar{d}$ π^+	$u\bar{s}$ K^+	$u\bar{c}$ $\bar{D}^0$	$u\bar{b}$ B^+
	$d\bar{d}$ π^0,η,η'	$d\bar{s}$ K^0	$d\bar{c}$ D^-	$d\bar{b}$ B^0
		$s\bar{s}$ η,η'	$s\bar{c}$ D_s^-	$s\bar{b}$ B_s^0
			$c\bar{c}$ J/ψ	$c\bar{b}$ B_c^+
				$b\bar{b}$ Υ

图 4.64　普通介子。图中强调显示的是最后一种发现的普通介子——B_c 介子。图片来源：费米实验室

而 ν_l 表示与其相应的中微子，该衰变过程的宽度相对于 B^+ 衰变的总宽度是 1.6%。理解这种特别衰变通道的一种方法，是把它看做在 B^+ 粒子内部环境下，b 反夸克通过弱相互作用转变为 c 反夸克加上一个正轻子和一个中微子的过程，人们所测量的正是这种可能过程。该衰变只是 100 多种 B^+ 粒子衰变通道中的一种，有些衰变通道已经被观测到并测量过了，还有一些通道人们已经开始寻找，但至今尚未观测到。

为了积累大量的观测数据，艰巨的任务落到了费米实验室的许多实验合作小组以及那些 SPEAR、DORIS Ⅱ 还有 CESR 电子-正电子对撞机研究小组的肩上。他们已经将详细描绘这些新发现粒子的性质以及它们之间的关系作为他们的本职工作。Tevatron 的质子-反质子对撞机可以很容易地生产 B 介子并测量其寿命和质量，再加上 LEP 和 SLC 的 Z 粒子衰变研究，共同带来了新的信息——Z 玻色子会发生 $b\bar{b}$ 衰变，这是一种弱相互作用衰变，在

其包含 b 夸克的产物中，B^+ 和 B^0 粒子各占 38%，B_s 占 11%，Λ_b 占 13%。在 1995 年英国曼彻斯特举行的一次会议上，LEP 小组公布了首次观测到另外两种底夸克重子——Σ_b 和 Ξ_b 的迹象。

研究 b 夸克的部门特别多，尽管在像 CESR 这种威力强大的机器上，可以以每年成百万的速度生产 B 介子，并且还能产生差不多数量的粲介子和 τ 轻子对，但物理学家们还是想得到更多的粒子。那只能去 B 粒子工厂。

工厂中的机器同加速器的设计走的是两条平行的路线。一条路线可能会被称作“发现”机，就像 Tevatron 这类机器，其目的是达到之前所未达到的更高能量，这样就能够进入一个未知的、等待人们探索的新物理领域。与之相反，工厂中的机器在一个较低的能量范围工作，目的是大量生产某种粒子。它们在能量上的劣势会通过其粒子束亮度和统计精确性来弥补。对于在工厂中工作的实验人员，新的物理内容来自于在可行的精度下探索和发现一些性质，比如特殊的衰变方式或者与理论不符的衰变可能等。这些都是那种必须通过积累上百万次的事件才能发现的东西，而不是一两次偶然的观测就能发现的。

在 SLAC，由 PEP 升级而来的机器——不对称 B 粒子工厂（图 4.65）属于新一代的工厂型机器。其非对称性来自于该机器是由两个具有不同能量的独立存储环组成的。原来的 PEP 对撞机被重新建造，能存储 9 吉电子伏的电子束，另外又新加了一个能存储超过 3 吉电子伏正电子的全新存储环，由 SLAC 线性加速器加上阻尼环来提供粒子束。支持这种不平衡对撞的伟大设计思想就是：能量更高的电子束会给碰撞碎片带来一个整体的单方向运动，因而更容易辨析发生的过程。

要建一座能完成这项任务的加速器并不那么容易，因此一定要存在很充分的理由，才能让机器的设计者们甘愿面对重重困难，来实现所需的功能。这样的充分理由确实存在。要大量生产

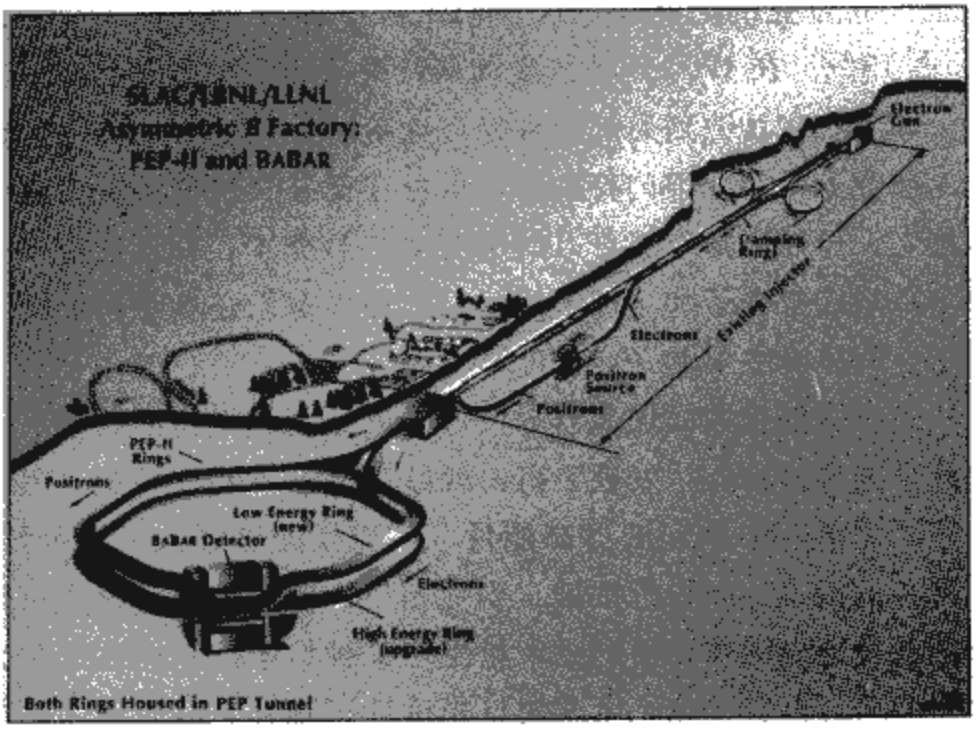


图 4.65　SLAC 的不对称 B 粒子工厂。注入器能够同时提供电子和正电子。图片来源：SLAC

粒子就表示要调整加速器，使其碰撞能量和要生产的粒子所需的能量完全一致——这是一种共振，在发生共振时碰撞截面特别大。由于不存在剩余能量转化为碰撞产物的动能，因此这些产生的粒子运动很慢。为了给它们提供动能、使它们运动更快一些，就要将加速器的能量上调一些，但是这时碰撞截面就会减小，因为加速能量不再共振了。为什么需要让碰撞产物运动得更快一些呢？因为从静止的探测器看上去，高速运动粒子的衰变速度，会比它们运动很慢或者根本不运动时更慢一些，这就是狭义相对论中的时间膨胀效应。衰变进行得越慢，就意味着静止时在万亿分之一秒之内就会完成衰变的 B 介子运动时的寿命会变得越长，这样探测器就会有更多的机会观测反应过程中到底发生了什么。为了在这两方面都做得很好——即在共振能量下碰撞的产率会特别

高，产生的高速粒子会产生时间膨胀效应，解决办法就是建造不对称加速器。

在 SLAC 不对称 B 粒子工厂中，用于观测那些漂移着的碰撞过程的机器，是一种新型的探测器，叫做巴伯（BaBar）。它是取了 $b\bar{b}$ 的谐音，听起来像是童话故事中那位大象英雄巴伯的名字，因此大象巴伯被用作该合作小组的标志①。探测器巴伯每年会探测到上亿个包含 b 夸克和 c 夸克的介子。

1999 年，和 SLAC 的机器一同启动的，是一台日本自己制造的工厂型机器，叫做 KEKB，即 KEK B 粒子工厂。它具有容纳 3.5 吉电子伏正电子和 8 吉电子伏电子的存储环，其不对称的能量同样用作给碰撞碎片施加冲击，这台机器利用新型的 BELLE 探测器进行监测，见图 4.66。图 4.67 表示的是由 BELLE 探测器所记录的一次 B 介子事件。研究 B 介子也是位于康奈尔的 CLEO 探测器 CLEOⅢ和 DESY 的 HERA-B 实验升级的主要目的。

所有这一切计划都将注意力瞄准了看起来很低调的 B 介子。B 介子之所以会吸引那么多的注意力，其中的一个原因就是，它能够澄清一个宇宙学中的重大问题：为什么在宇宙中占统治地位的是物质而不是反物质？答案可能就在于，桀骜不驯的 B 介子拒绝遵守一种称作 CP 对称性的联合对称性。

C 对称性联系着粒子和它们的反粒子。称作电荷共轭（charge conjugation）的 C 操作表示将每个粒子都替换成它的反粒子，反之亦然。如果 C 是一种有效的对称性，那么所有东西都会仍像以前那样运转，强相互作用和电磁相互作用都遵从这种对

---

① 译者注：大象巴伯是全法国儿童最喜欢的卡通形象，它是一位生活在丛林之中的大象国王。该形象第一次出现在法国作家 Jean de Brunhoff 的童话故事书《巴伯的故事》（1931 年）中，接下来他又写了另外六本关于巴伯的书。Brunhoff 死后，他儿子继续出版了一系列有关巴伯的故事书，后来又出现许多描写巴伯的影视作品。大象巴伯的形象在全世界开始广为流传。

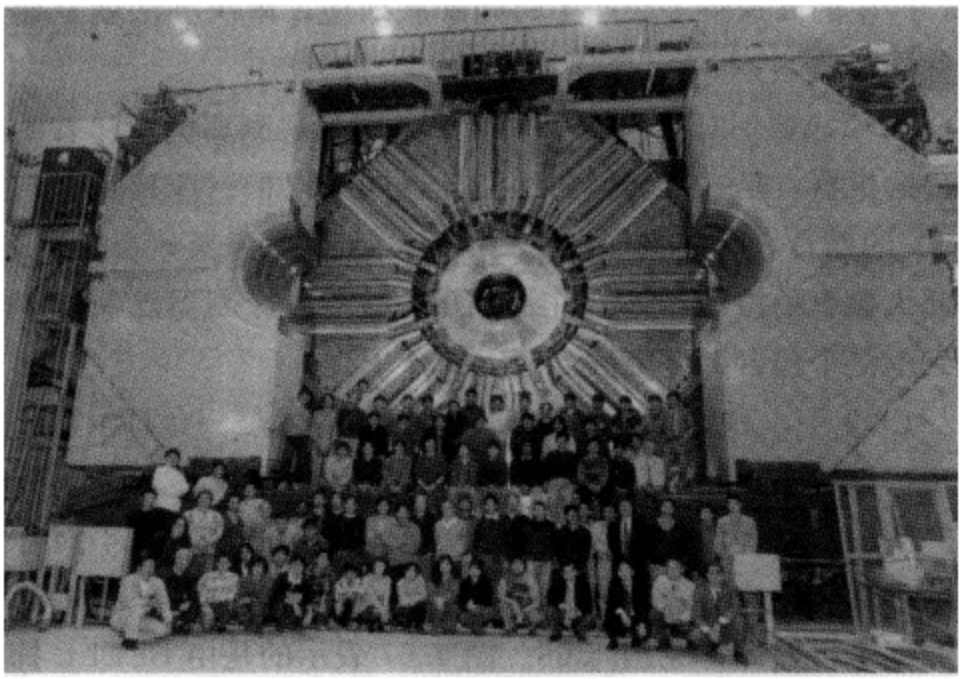


图 4.66 BELLE 探测器的巨型闸门。位于 KEKB 的 BELLE 探测器马上就要“起航”了。加速器束管道穿过了中间的大洞。“闸门”是 μ 子- K子探测器末端的盖子，当它关闭时能盖住大多数探测器。图片来源：KEK-BELLE 合作小组

称性。它们还遵从 P 对称性（奇偶对称性，P 为 parity 的缩写），它也被称为镜像对称性。P 操作是指将每一件东西都取其在平面镜中的像，例如将两个分别为左旋和右旋的自旋互相对调，任何东西都不发生任何改变。但对于弱相互作用来讲，C 和 P 都不是有效的对称性，中微子可以清楚地体现这个特征。左旋的中微子和右旋的反中微子占绝大多数，因此它不遵从 C 对称性，因为 C 对称性表示，将一个左旋的中微子换成左旋的反中微子；它也不遵从 P 对称性，因为在 P 操作下，一个左旋的中微子将变成一个右旋的中微子。但是如果将这两种对称性放到一起，C 和 P 的组合将会把左旋的中微子和右旋的反中微子联系在一起，一切就没有什么问题了。

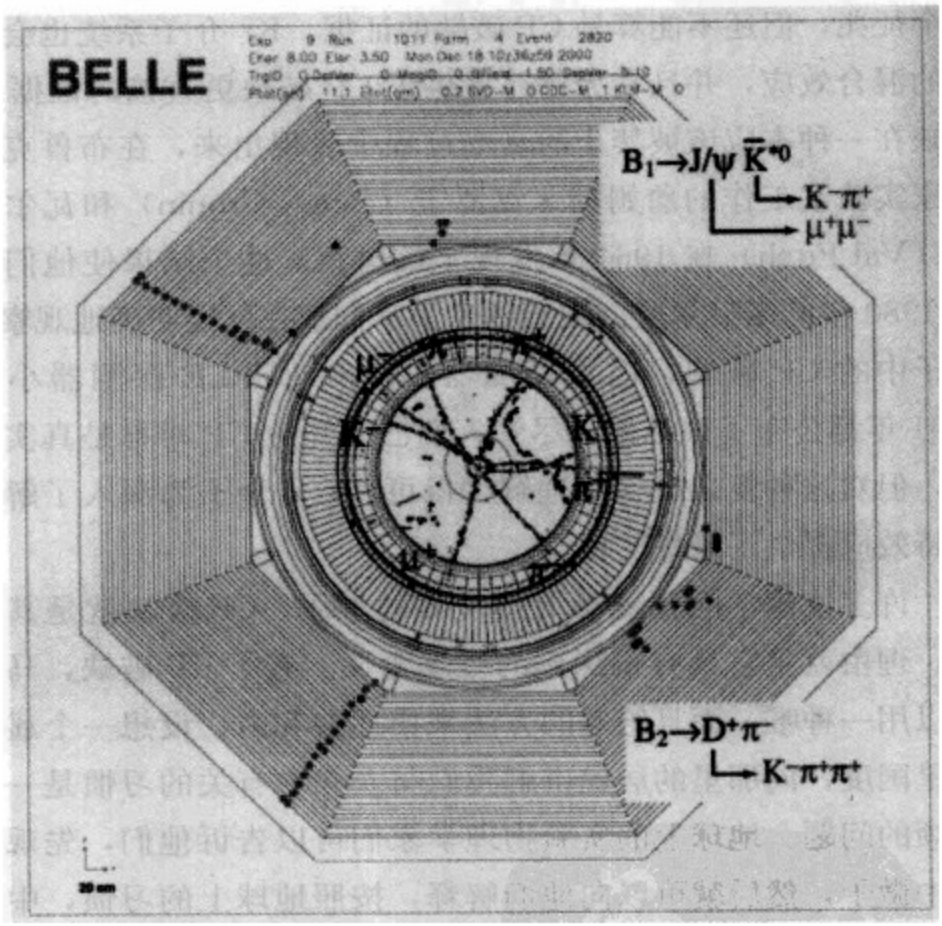


图 4.67 BELLE 探测俘获到的两个正在衰变的 B 介子，每个 B 介子都差不多有一整个氦原子那么重。两幅插图显示了各自的衰变路径。图片来源：KEK-BELLE 合作小组

大自然又给了粒子物理学家们一个烦人的问题。在一种混合（mixing）效应中，B^0 介子和它的反粒子 $\bar{B}^0$ 可以互相转化。换句话说，一束由纯 B^0 介子所组成的粒子束向前运动，一会儿粒子束就会变为由 B^0 介子和 $\bar{B}^0$ 混合而成的了，即使你不去碰粒子束，这种变化还是会发生。这种 B 介子混合效应是由 ARGUS 小组（A Russian-German-US-Swedish collaboration，俄罗斯-德国-美国-瑞典合作小组）于 1987 年利用 DORIS Ⅱ 电子-正电子对撞机首次发现的。这是 CP 对称性破缺（或者简称为 CP 破缺）的

一个征兆，但还不能算是 CP 破缺的证据。K^0 介子系统也会有类似的混合效应，并且它另外还提供了 CP 破缺更直接的证据。CP 破缺在一种本应该被禁止的衰变过程中表现出来，在布鲁克海文国家实验室工作的詹姆斯·克罗宁（James Cronin）和瓦尔·菲奇（Val Fitch）在 1964 年发现了该过程，这个结果使他们获得了 1980 年的诺贝尔奖。实验物理学家们现在已经明确地观察到 B 介子中的 CP 破缺，巴伯探测器小组和 BELLE 探测器小组在 2001 年都公布了该结果。尽管人们已经知道 CP 对称是真实存在的，但对这种复杂过程的理解少得可怜，B 介子为深入了解这种奇特效应提供了最好的机会。

许多物理学问题能够揭露世界的本质，CP 破缺就是其中之一。理由就是它具有很深刻的哲学含义。有了 CP 破缺，马上就可以用一种唯一而且绝对的方法来描述左和右。设想一个遥远的外星国度，向那里的居民讲解我们与左和右有关的习惯是一个很麻烦的问题。地球上的实验物理学家们可以告诉他们，先观测一下中微子，然后就可以向他们解释，按照地球上的习惯，中微子是一个与左手有关的东西。啊哈，收到外星人的回复了："我们怎么知道哪一种是中微子，哪一种是反中微子?"这个问题的答案必须用一些 CP 破缺过程来表达，例如可以比较 K 介子的衰变过程，它要么衰变成 $\pi^- + e^+ + \nu_e$，要么在极少的情况下衰变为 $\pi^+ + e^- + \bar{\nu}_e$。

通过 CP 破缺，K 介子和 B 介子衰变可以用来区别物质和反物质，这令宇宙学家们非常激动。在他们所知的宇宙范围内，物质要比反物质多很多，占宇宙中物质的绝大多数。如何解释这种非对称性，被看做是宇宙学家们理解我们的宇宙以及探求宇宙产生问题所面临的一种根本挑战——在宇宙诞生的大爆炸假定中，不存在任何与之相关的线索。人工加入一些不平衡性也不是一种令人信服的解决方案。但很有可能早期宇宙中出现的 CP 破缺效

应，导致了我们今天所看到的物质大量失衡。

这种由CP破缺所带来的不可思议的事情还有许多。人们发现四夸克标准模型不符合CP破缺效应的要求——四种夸克所提供的自由度不足以容纳CP破缺。早在1973年——甚至比第四种夸克粲夸克的发现还要早一年，日本东京大学的小林诚（Makoto Kobayashi）和盖川敏英（Toshihide Maskawa）就意识到，在四夸克模型的图像上再加上几种夸克，是允许这种自由度存在以及将CP破缺纳入现在所谓的标准模型中的唯一方法。根据小林诚和益川敏英的想法，再假定夸克都是成对出现的，粒子物理学家们推测一共存在6种夸克。B粒子工厂的灵敏度特别高，足以检验小林诚和益川敏英提出的存在更多夸克的理论以及由此而引起的任何修正。肯定会存在某些变动的。因此根据实际情况，B粒子工厂以一种相对较低的能量，来寻找存在更多夸克的间接证据。但实验物理学家们现在已经有了很多的直接证据，证明还存在另外一种夸克：顶夸克（top quark）。

顶夸克

《纽约时报》1994年4月26日的头版用很长的篇幅写道：

> “由古希腊哲学家提出的理解物质本性的问题，可能已经在伊利诺伊州的巴特维亚，随着顶夸克证据的发现而解决了。就现在所知，顶夸克是组成整个物质世界的12种基本砖块中的最后一种。”

这是一篇夸张的文章。它所报道的是CDF合作小组公布的、在费米实验室的Tevatron上进行的质子-反质子对撞过程中所出现的12次顶夸克事件。第二天，伦敦的《泰晤士报》以“海外新闻”为标题，刊登了一篇报道该发现深刻意义的文章，文中

写道：

"如果得到了证实，它将完成对标准模型的证明，标准模型是物理学家们现在对原子如何构成这一问题的回答。"

CDF 小组在 4 月 22 日提交了一篇 152 页的论文预印本，公布了他们的结果。在随后的新闻发布会上，他们对全世界的媒体讲解了他们结果。他们指出，由于没有找到足够数量的可能顶夸克事件，因此不能排除该结果是由某些统计偏差引起的。那时，他们看到的疑似顶夸克的事件，同背景事件的比例仅仅为 1∶400。此外，在同一场新闻发布会上，他们的竞争者 D0 合作小组（他们的探测器也在 Tevatron 环上，距 CDF 有大约 1/3 周长的位置）报告称，他们还未搜集到任何有关顶夸克的证据。随着时间的推移，人们继续对数据进行分析，积累数据。在第二年——1995 年，两个小组都提供了清晰而明确的证据。他们证明了顶夸克的存在。顶夸克——第六种、也可能是最后一种夸克——被发现了。

顶夸克的质量大得不可思议，为 174.3±5.1 吉电子伏，几乎和一个完整的金原子一样重。这个巨大的数字需要让我们考虑一会儿：一种基本粒子竟然和一个较大的原子一样重！正是如此巨大的质量，才解释了为什么只有费米实验室才能直接看到顶夸克——因为只有费米实验室具有足够的能量。

根据 QCD 理论，在费米实验室产生的大约 90%的顶夸克-反夸克对，都是通由最初的质子和反质子所形成的夸克-反夸克对发生湮灭所产生的，见图 4.68 的图（a）。剩下的 10%来自于胶子同胶子的反应，见图 4.68 的图（b）至图（d）。顶夸克最可能通过弱相互作用衰变为一个 b 夸克和一个 W 玻色子，其过程为 $t \rightarrow W^{+} + b$，反顶夸克的衰变过程则是通过 $\bar{t} \rightarrow W^{-} + \bar{b}$ 实现的。

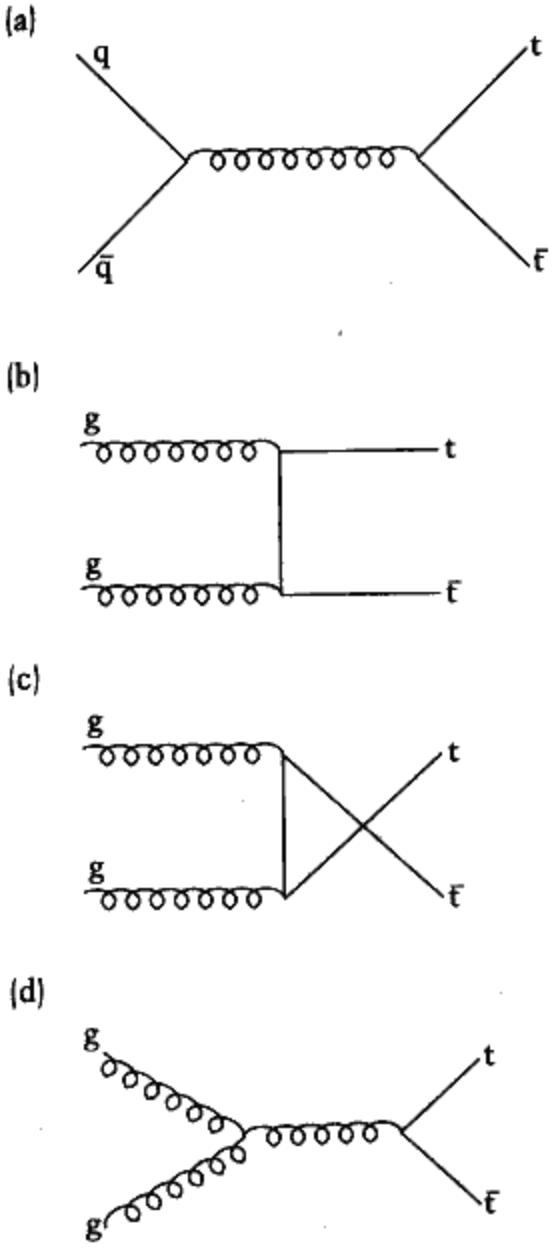


图 4.68 由质子-反质子碰撞产生顶夸克-反顶夸克过程的最简单费曼图。图(a) 为夸克-反夸克碰撞，图(b) 至图(d) 为胶子-反胶子反应

接着，W 粒子可能会生成一个轻子（一个电子、μ 子或 τ 子）加上一个相应的中微子，即 $W^+ \to 1^+ + \nu_l$ 和 $W \to 1^- + \bar{\nu}_l$，其中 1 表示轻子。另外，W 还可能衰变为夸克-反夸克对，即 $W \to q + \bar{q}$。

顶夸克不仅能成对产生，也可以单独产生，例如通过q+b→q+t的过程产生，但是在实际中该过程发生的机会特别小。

无论产生的机理是什么，顶夸克产生的可能性都非常的小。CDF 小组和 D0 小组 8 年间的数据中只积累了不到 200 个左右的顶夸克对事件。CDF 小组推断其产生的碰撞截面是 7.5±1.9 皮靶恩，D0 小组给出了更低碰撞能量下 5.2±1.8 皮靶恩的反应截面值。换句话来说，碰撞每产生大约 300 万个 $\bar{b}b$ 对才能产生一对 $\bar{t}t$。或者再换一种说法，在 Tevatron 的能量下，大约每发生 100 亿次质子-反质子碰撞过程，才能产生 1 个顶夸克-反夸克对。

随着实验者们不断积累更多的数据，这种散射截面的测量值也越来越精确，人们可以用它来很好地检验 QCD 理论。现在能够得到一个精度不是很低的计算值，利用微扰 QCD 理论进行计算，可以得出在 Tevatron 上 $\bar{t}t$ 产生的碰撞截面值为 4.8 皮靶恩，在测量精度范围之内与实验相符。理论物理学家们相信重夸克计算的微扰理论。这是因为一个重夸克（没有比顶夸克还重的夸克了）将计算的能量尺度，置于一个强相互作用很小的范围之内，因此微扰论就有意义，这一点将会在第五章再次谈到。使得这种独特的计算更值得信赖的另一个因素，就是大多数 $\bar{t}t$ 对都由来自于质子-反质子碰撞的夸克和反夸克产生。质子和反质子是自由相互作用的夸克和反夸克的源头，理论物理学家们对他们的这一理论十分满意，比他们用胶子做的类似描述还要满意。也还有一些因素是不确定的，比如强耦合参数精确值的使用等。还存在一个棘手的问题，这与随着顶夸克的产生，探测器中观测到的大量轨迹有关。可能还存在一些科学上未知的衰变过程，但理论物理学家们在估算的时候没有将它们包括进去。如果它们真的存在，那么在将来实验物理学家们能够测量各种各样的顶夸克衰变通道时，这些衰变过程就会被揭示出来。

寻找顶夸克的人还肩负着另一项任务。顶夸克的质量加上 W

玻色子的质量，就可以对高能物理史上最受欢迎的粒子——西格斯粒子（Higgs particle）的质量加以限制。这是因为 W 粒子一生中的部分时间是和 Higgs 粒子及顶夸克结合在一起的，这由它的量子本性决定。这就说明，三种粒子的质量存在一定关系。Higgs 粒子同物理学家们至今所观测到的任何粒子都不相同，它参与了粒子间相互作用的对称性破缺，并且是解释粒子怎样获得质量的理论体系的一部分。在有关弱相互作用的电弱理论中，对称性破缺和 Higgs 粒子是正确预言 W 粒子和 Z 粒子质量的核心。很明显，Higgs 粒子不是一个小人物。

然而 Higgs 粒子以及它的对称性破缺方案还有一些问题。如果真的找到 Higgs 粒子，将为至今仍然有点神秘的标准模型砌上最后一块坚实的砖块。在此过程中，顶夸克扮演了一个特殊的角色，因为夸克- Higgs 粒子相互作用的强度同夸克的质量成正比，顶夸克具有特别大的质量，这就表示顶夸克比其他任何一种夸克都更容易同 Higgs 粒子相互作用。这一点使得在 Higgs 粒子的产生和衰变过程中，顶夸克比其他各种味道的夸克都要重要得多。此外，被人们所接受的对称性破缺理论认为，顶夸克到 W 玻色子的衰变过程有特殊的角分布，这是因为人们预计顶夸克会生成极化 W 玻色子的特定混合。将来的实验数据会让实验者们观测到这种效应。

到目前为止，实验物理学家们还没有发现 Higgs 粒子。在 LEP 上进行的电子-正电子实验显示，Higgs 粒子的质量一定超过 65 吉电子伏。理论物理学家们预测它的质量应该为100～1000 吉电子伏。出现这样大的质量范围，是由于受现在测得的 W 粒子和顶夸克的质量精度限制，以及 Higgs 粒子的质量对其他两个质量的依赖性很弱。如果 Higgs 粒子不能马上现身的话，研究者们都期盼能在 CERN 的 LHC 上发现它。撇开 Higgs 粒子不说，LHC 生产顶夸克的能力要比 Tevatron 强，大约每秒钟能产生 7

个 $\bar{t}t$ 对。

从许多方面来看，发现顶夸克都不是一个意外。从 1977 年发现底夸克开始，粒子物理学家们就相信，在某个地方存在着它的搭档。实验物理学家们都主动地寻找顶夸克，由于他们在低于 Tevatron 能量的条件下没有发现顶夸克，他们逐渐将它的质量下限推得越来越高。即使在最后测出顶夸克的质量时，人们也没有感到特别吃惊。在 1995 年，一个 LEP 工作小组反过来研究了 Higgs 粒子、W 粒子和顶夸克之间的质量关系，通过合理猜测 Higgs 粒子的质量，计算出一个顶夸克的质量值。他们预言顶夸克的质量在 150～203 吉电子伏，最可能的质量是 178 吉电子伏，这个值同测量值非常接近。

在费米实验室发现顶夸克，标志着经历 17 年的搜寻顶夸克的工作结束了，但这次结束本身又是一个全新的开始。找到顶夸克无疑是一项令人激动的发现。如果没找到它，肯定也将会是一次令人激动的失败。

大草原下的原子粉碎机

发现顶夸克的那台机器——Tevatron——是世界上最强大的粒子加速器。它位于伊利诺伊州巴特维亚的费米实验室，在芝加哥以西 60 千米。但是 Tevatron 自己却不能生产顶夸克。它要靠能量较低的机器为它提供。实际上，Tevatron 是一条流水线中的最后一台加速器，这条流水线是由协同工作的 5 台加速器共同组成的，见图 4.69 至图 4.71。

整个实验起始于一个立方形金属盒——离子源——的内部，盒子中的氢原子都被加上了一个电子，形成了一个由一个质子和两个电子所组成的负离子。然后这些离子通过一个静电加速器 75 万伏的电压进行加速，这种加速器同考克饶夫和沃尔顿在 1932

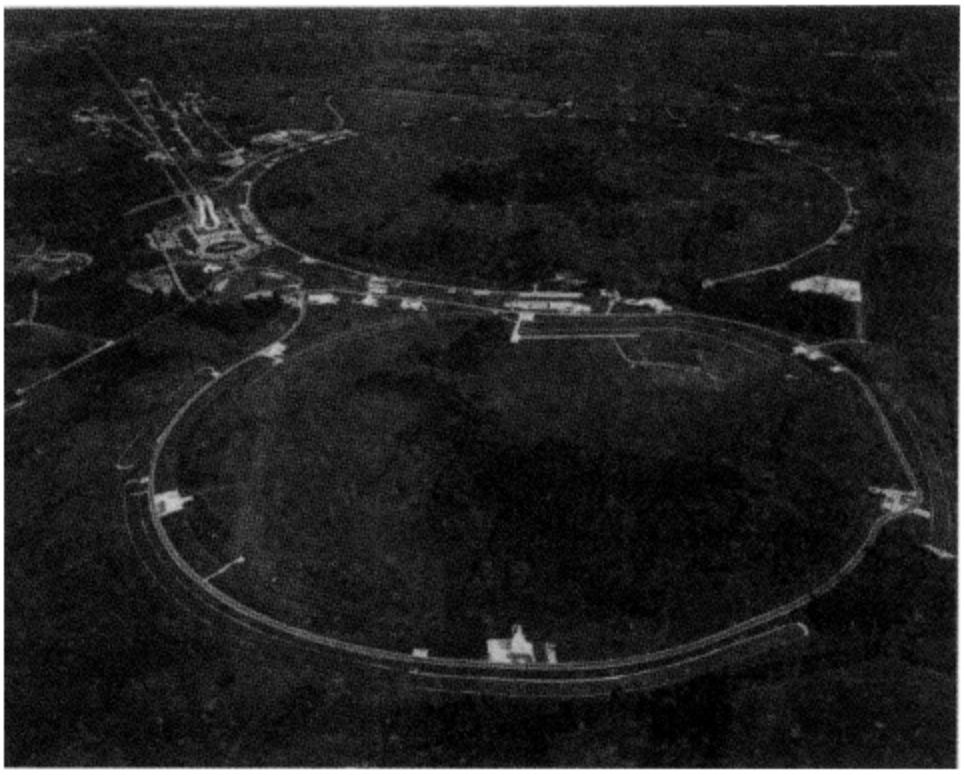


图 4.69 费米实验室加速器的鸟瞰图。前面的环是主注入器。朝后的环是 Tevatron 本身。费米实验室的其他仪器，都靠在两环衔接瓶颈处的左侧。图片来源：费米实验室

年用来击碎原子的那台加速器属于同一种类型。第二个加速阶段是在一台长 150 米的线性加速器上将离子的能量提高到 400 兆电子伏。接着这些离子通过一层很薄的碳，它能够将离子上的电子夺走，只留下质子进入第三个加速阶段。这是一台周长为 500 米长的“小型”质子同步加速器，叫做辅助放大器（booster），它能将质子的能量加速到 8 吉电子伏。

以上所说的都仅仅是进入大型加速器前的准备阶段。从辅助放大器出来后，质子就进入了主加速环（main ring），它是一台周长为 6.28 千米的旧质子同步加速器，能够将质子的能量提升到 150 吉电子伏。最后，质子就要遇到强大的 Tevatron 了。它之

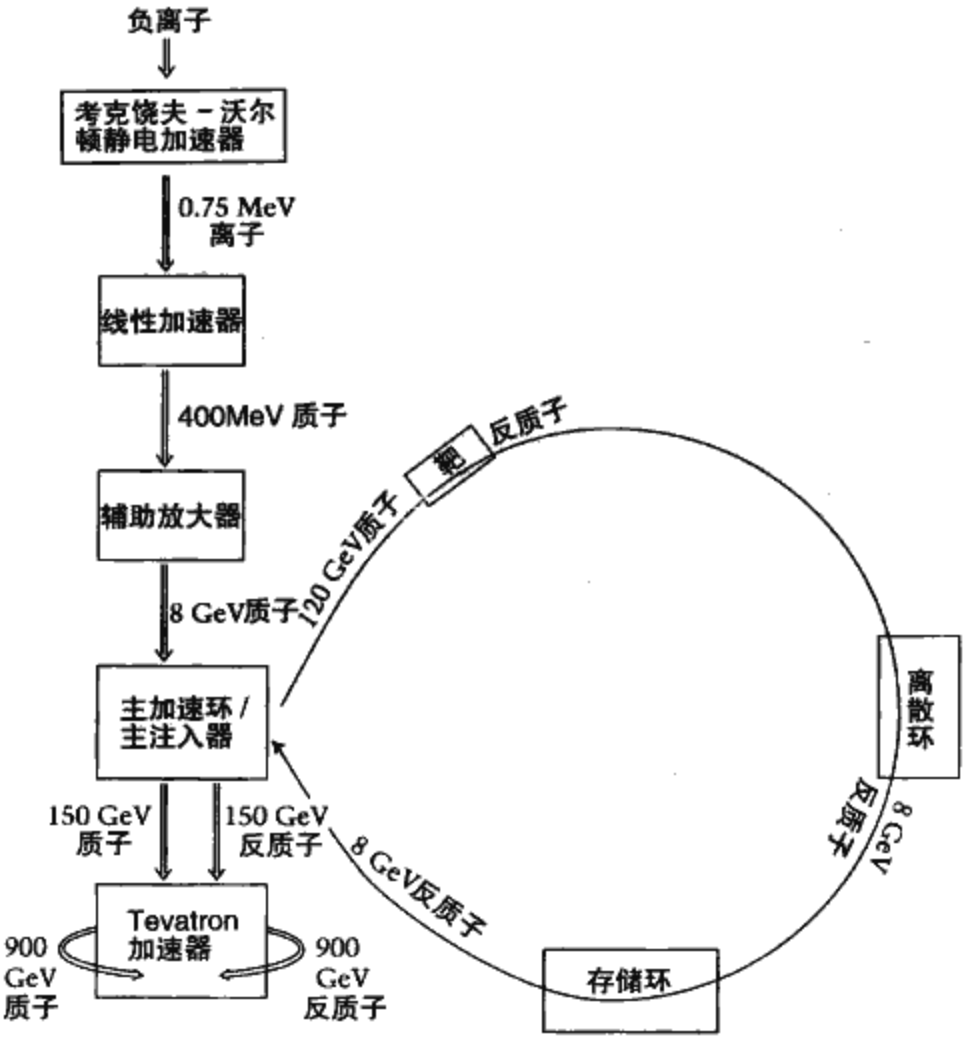


图 4.70　说明 Tevatron 对撞机主要特征的流程图。主注入器将旧的主加速环代替了

所以被称为 Tevatron，是因为它能够将质子能量最高加速到 1 太电子伏（tera electronvolt），也就是 1000 吉电子伏，它在 1984 年投入使用，是世界上第一台超导同步加速器。它就安放在主加速环的正下方，与主加速环共用一个隧道。它有 1000 多块超导磁铁，其数量和主加速环上的普通磁铁差不多，这些超导磁铁能够在比较小的能量消耗下提供更强的磁场。它能够使质子和反质子

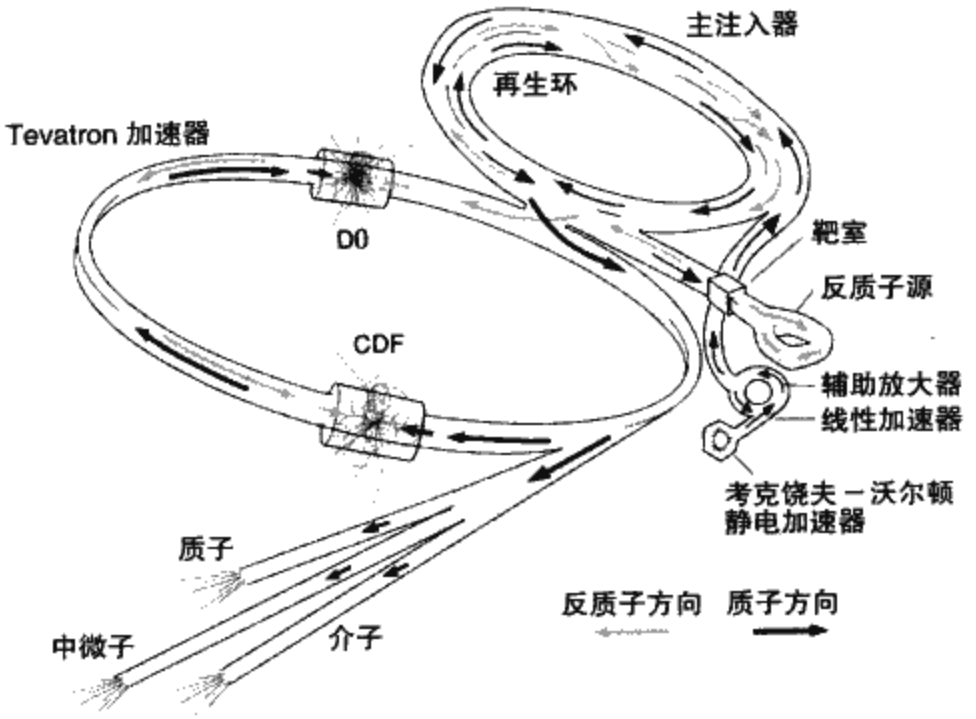


图 4.71 Tevatron 加速器链的草图。图片来源：费米实验室

在两个点发生总能量为 1.8 太电子伏的对撞（或者说每一束粒子的能量为 900 吉电子伏），并利用两个探测器 CDF 和 D0 进行观测。

但我们有一些东西没说，那就是反质子。产生质子是很容易的——只要将氢原子上的电子剥走就得到了质子，但产生反质子还是有一些麻烦的。实验者们通过将一束质子射入金属靶来产生反质子，反质子是通过金属原子核内部的核子相互作用产生的。问题是这些反质子以一种混乱的粒子束的形式出现，而不是一束有序的粒子束，因此不能够把它直接注入同步加速器。

为了驯服它们，工程师们设计了一套处理粒子束的巧妙方法，称作随机冷却，这是由西蒙·范·德·梅尔（Simon van der Meer）在 1972 年发明的。该思想在本质上就是让反质子在存储环中运动，在环上的某些点用电传感器检测粒子束相对于圆环中

心轴线的分布情况。该信息很快就传递给位于反质子前面的一个电“打击器”上。打击器会令这些粒子偏转到正确轨道上来。随着这种过程一圈一圈地进行，反质子就迅速变成了一束理想的粒子束，然后被存储起来。越来越多的反质子不断添加进去，直到聚集了足够多的反质子进入下一步的加速程序。“随机冷却”的“随机”是指粒子束取样的任意性，“冷却”是指降低反质子扩散速度，这同降低气体分子运动速度以使其温度变得更低的过程相类似。随机冷却是推动实验粒子物理学向前发展的一种伟大而可行的方法。这种方法第一次在 CERN 的 ISR 进行试用，它被应用于 $Sp\bar{p}S$，使鲁比亚和他的小组找到了 W 玻色子和 Z 玻色子。实际上，范·德·梅尔正是由于他在随机冷却方面的工作，同卡罗·鲁比亚（Carlo Rubbia）一起分享了 1984 年的诺贝尔物理学奖。

随机冷却在费米实验室中发挥了重要作用。反质子首先通过将从主加速环导出的 120 吉电子伏的质子束打到铜靶上制造出来，它们从那里再进入周长为 520 米的 Debuncher 环进行冷却。经过冷却的反质子接下来被存入存储环当中，直到聚集足够的数量。包括随机冷却在内的各种冷却方法施加在到处乱闯的反质子上，使粒子流的密度增加了 100 多万倍。最后，冷却的反质子被注入主加速环中，在到达 Tevatron 之前，它们的能量在主加速环中被加速到了 150 吉电子伏，形成了与早已等在那里的质子相对称的反向运动粒子束。

至少这就是在 CDF 小组和 D0 小组发现顶夸克时，所有环节的运行方式。在那以后，Tevatron 又经过了升级。升级的中心环节是兴建了一台全新的加速器——一台叫做主注入器的质子同步加速器。主注入器的安放位置紧靠在 Tevatron 的上方，其周长为 3.3 千米，大约是 Tevatron 的一半，它的任务就是代替主加速环。实际上，主加速环的一些部件已经被用到了新机器上。主注

入器耗资约2.6亿美元，于1999年投入使用，建造它的主要原因是想消除由主加速环引起的瓶颈效应——主加速环无法输送足够多的粒子。150吉电子伏的主注入器依靠更加强大的聚焦能力，能够为Tevatron提供密度更大的粒子束。它能够将粒子束的亮度提高5倍。其他的改进之处就是增加了另一个环——8吉电子伏的再生环，它同主注入器共用一个全新的隧道，它的功能是存储经过Teratron碰撞周期之后剩下的那些反质子。总的来说，这些改进方案使Tevatron的整体亮度提高了大约10倍。

这次改进计划使得费米实验室能够更好地参与一整套改进后的新实验，比如寻找“中微子振荡”（在中微子震荡过程中，一种类型的中微子能够转化成另一种类型的中微子），再比如研究B介子的性质以及更近距离地寻找违反标准模型的现象。其中肯定能从中受益的一个计划就是顶夸克性质的研究。新的Tevatron亮度更高，总能量也得到了一些提升——增加到了2太电子伏，其顶夸克对的产生率也比旧机器提高了20倍。旧机器只能成百个地生产顶夸克，而在2001年启动的“Run Ⅱ”——改进后的Tevatron——则可以成千地生产。

发现顶夸克

Tevatron并不是唯一一台经过翻新的仪器。CDF探测器以及环对面的D0探测器都经过了升级。这种升级是必需的，因为经过翻新的Tevatron亮度更高，会产生特别多的碰撞，产生的信号能够将旧探测器淹没。

CDF探测器是两个探测器中年龄较大的一个，它是最典型的现代粒子物理探测器，与它有关的数据令人吃惊。CDF合作小组大约有600人，其成员来自于美国及全世界的30多个研究所，它由两位民主选出的发言人以及其他官员和执行委员会所领导。现

在我们可以公布以下这台机器的完全数据：CDF 探测器重约 5000 吨，17 米长，宽和高都是 12 米——其大小足以装下一个普通家庭的房子。

CDF 探测器是一台由几个不同部分组成的通用探测器。因为它观测的是两束能量相同的粒子束发生的迎面碰撞，碰撞产物会散布得到处都是，因此那些不同的探测器部件都被制成了圆柱的形式，这样就可以包围尽可能多的碰撞点，这些探测器看起来就像是粒子束路径上的一段管道。这同固定目标实验的探测器设计恰好相反，固定目标探测器的任务是寻找投射在粒子束入射方向一定角度之内的反应产物。像 CDF 这样的探测器很像一套俄罗斯套娃，每一种探测元件都只完成一种测量任务，位于其四周的元件完成别的任务，而这些元件周围又环绕着其他元件，如此反复下去。但从功能的角度来看，可以很清楚地将这些元件划分为两类。如果 CDF 探测器是一只苹果，它的核就是轨迹测量系统。这个核由圆柱形的磁铁组成，粒子束沿着圆柱的轴运动，环绕着粒子束的探测器可以精确地测量出带电粒子的位置。苹果核外面的果肉是测量粒子能量的“量热器”系统，见图 4.72。

在翻新后的 Tevatron 中，用来观测碰撞的是升级后的 CDF 探测器，即 CDF Ⅱ，也就是第一个看到顶夸克的那台探测器。它最里面的一层最接近于碰撞点，由全新类型的 CDF 硅顶点探测器组成。它是一个由 5 层同心的硅片层包裹而成的 1 米长圆柱，这 5 层硅片的厚度从 2.4 厘米到 10.6 厘米不等。从外边看，它就像是一个齐腰高的排水管。其基本探测元件是硅微带探测器（silicon microstrip detector），它利用了半导体（也是晶体管的组成物质）的物理性质，可以将穿过的带电粒子的位置精确测定到 10 微米量级。其中，一个硅顶点探测器的主要任务是定位粒子路径中的分叉，或者叫顶点，这种分叉是由相互作用点附近的粒子产生或衰变所造成的。

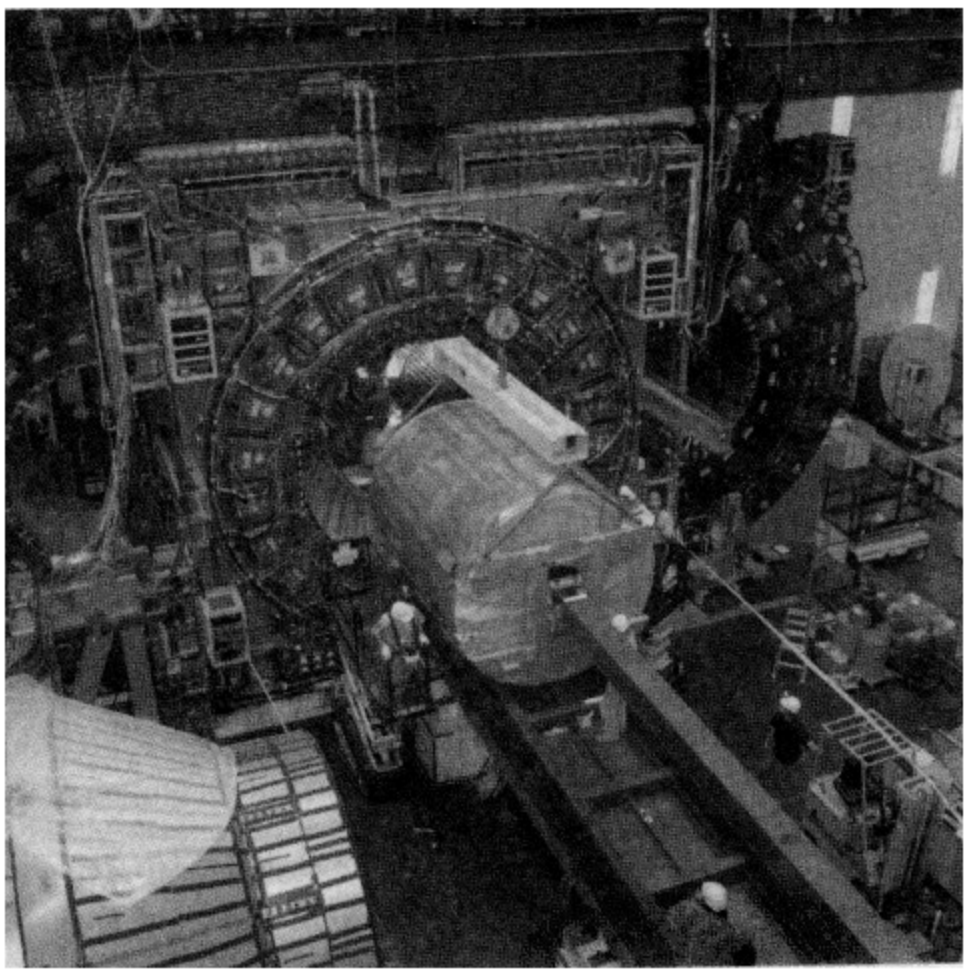


图 4.72　费米实验室的 CDF 探测器。中间的轨迹探测室正在安装。
图片来源：费米实验室

如果没有硅微带探测器，寻找顶夸克的工作将会变得更加艰难，它的精度比排在它之后、第二精确的电子探测器精确了 10 倍。如果没有理论物理学家给实验物理学家提供一些寻找顶夸克的线索，寻找顶夸克的工作可能也会变得更加困难。顶夸克的质量要大于 85 吉电子伏——实验物理学家们直到 20 世纪 90 年代才证明了这一点，因为 W 粒子的质量大约是 80 吉电子伏，b 夸克的质量约为 4 吉电子伏，顶夸克能够衰变为$t \to W + b$。这种衰变

方式非常快，其他味道的夸克不会以这种方式发生衰变，因为那些夸克都比 W 粒子轻得多。衰变快是因为顶夸克能产生真实的、物理上的 W 玻色子，而其他像 b 夸克和 c 夸克这样的“重”夸克衰变只涉及虚 W 玻色子，见图 4.73。因此其他夸克在通过虚 W 粒子发生衰变之前，必须等待合适的量子能量起伏，而顶夸克则能够勇往直前，不需要任何耽搁就能发生衰变。它衰变的时间大约是 10^{-24}秒。与之相反，b 夸克和 c 夸克的寿命要比它长 1 万亿倍，其寿命大约为 10^{-12}秒。

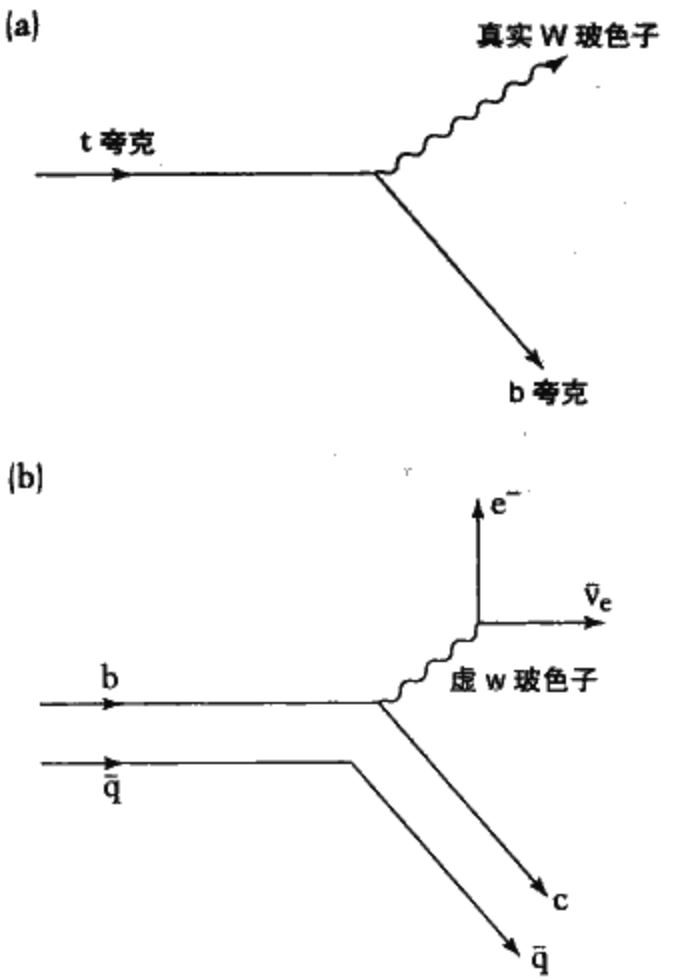


图 4.73　(a) 顶夸克可以衰变为一个真实的 W 玻色子和一个 b 夸克。(b) 图中作为 B 介子的一部分，b 夸克通过虚 W 玻色子进行衰变

多么可怜的顶夸克呀！它衰变得太快了，以至于一个新出炉的顶夸克都没有时间变成介子（而一个新产生的b夸克能够变成B介子）。顶夸克特别不稳定，因此它永远不可能形成普通的介子和重子，这就是为什么尽管存在顶夸克，在1998年发现的B_c粒子还是被宣称为最后一种被发现的常规介子。顶夸克是很特别的。其他所有的物质粒子，要么是寿命很长、很稳定的，要么就变成了复合粒子的一部分，而顶夸克却注定要过着短暂而孤独的生活。

我们先抛开伤感的情绪，顶夸克独特的性质为实验者寻找它们提供了线索。顶夸克的寿命如此之短，因此我们无法分辨顶夸克的衰变点和最初质子-反质子碰撞的相互作用点。W粒子本身的寿命和顶夸克的寿命差不多，因此由顶夸克衰变产生的W玻色子的衰变，看起来也像是在最初的作用点发生的。但是底夸克的衰变相对较慢，它在绝望地变成其他更轻的夸克之前，会运动约半毫米的距离。这就意味着在与一个底夸克或反底夸克相联系的粒子喷注过程中，会存在一个标志其衰变的顶点。因此在原始碰撞点附近出现的一个具有顶点的粒子喷注，很可能就是b夸克喷注。实际上，实验物理学家们都在积极地寻找这种性质，这是一个被他们称为标签（tagging）的例子——通过粒子或者粒子产生的性质来寻找某个特定过程。识别b夸克喷注和$\bar{b}$夸克喷注很重要：在CDF探测器中看到的大多数顶夸克事件，都包括由单个W玻色子衰变成一个夸克和反夸克对时形成的两个喷注和另外两个喷注（分别由t和$\bar{t}$衰变生成的b夸克和$\bar{b}$夸克产生），还有一个正电子或某种其他轻子（由另一个W粒子衰变而来，与轻子相应的中微子看不见）。

探测器的第一部分是硅顶点探测器，它记录了粒子的这种星状放射，并将其分解为表明顶夸克存在的特定喷注。从理论上来讲，至少它的基本单元——硅微带探测器是一种简单的技术。将

一层硅仔细地掺杂上一层可控层（例如磷），再将另一种材料（例如硼）的细条平行地植入。这些细条就是微带（microstripe），只有几微米宽，一般从一个细条的中心到另一个与之相邻的细条中心之间的距离是 50 微米。最后还要在硅基底以及细条上连接上金属电极，见图 4.74。

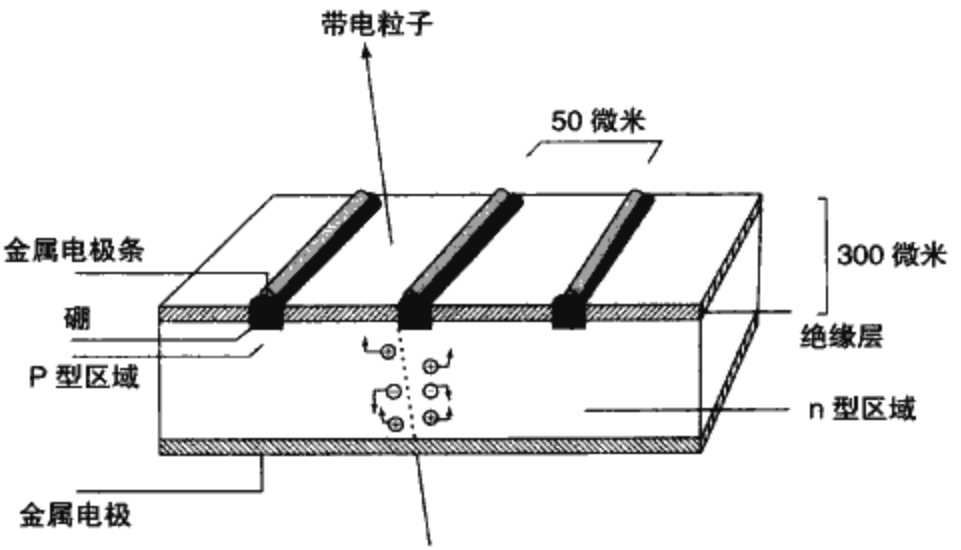


图 4.74　硅微带探测器。一个单独的带电粒子经过时，会在半导体材料中产生电荷，并通过电传导带产生可探测的电脉冲

在硅晶体中，1 个原子和它周围的 4 个原子相结合。磷具有 5 个可用的价电子，因此当它努力想让自己置身于硅晶体中的时候，其中的一个电子就剩余出来，很容易变成自由电子。这种自由电子能够携带电荷在材料中四处运动，因为它带负电，因此这种掺杂被称为 n 型掺杂①。硼与硅正好相反。它只有 3 个可用的价电子，因此那些渗入硅中的硼原子在应该存在电子的地方形成了一个空隙。这种空隙被称作空穴，它实际上少了一个带负电荷

① 译者注：n 是 negative 的缩写，表示“负”。

的电子，它的行为就像是一个运动着的正电荷，因此这叫 p 型掺杂①。在两者之间，两种掺杂会在界面上形成一个电场。

当一个带电粒子——比如说一个带电介子——通过这一界面，它能够将一些带电的电子和空穴从晶体中赶出来，从而留下一条它走过的轨迹。这种在电场作用下的脱离现象是由杂质原子决定的，空穴会朝着 p 型细带运动而电子会离开 p 型细带。结果会在一亿分之二秒的时间内产生一个电脉冲。一个经过的带电粒子会在几条细带上产生电流脉冲，这样就可以计算出它的位置和轨道。硅探测器极高的分辨率可以将像粲介子和底介子这种粒子的原始顶点（或称生成顶点）和次级顶点（或称衰变顶点）区分开来。通过顶点之间的间距，可以计算出粒子的寿命。

硅微带探测器在 20 世纪 80 年代发展迅猛，并立即运用到固定靶（fixing-target）实验中，它能够精确追踪当时新出现的粲粒子。第一次利用这种探测器技术观测的粒子束碰撞实验，于 1990 年在 SLAC 的 MARK Ⅱ 探测器上进行。第一次利用这种技术观测的强子对撞实验 1992 年在 CDF 探测器上进行。升级后的 CDF 探测器拥有一个第二代硅顶点探测器，这种设计是为了处理 Tevatron 升级后出现的更高的质子-反质子碰撞率。它测量一个事件时会激发比升级前多 10 倍的电信号，总计激发的电信号个数超过 40 万个。等待的电子器件必须在 10 微秒之内将这么多数据全部处理掉。

另外两个硅层位于距粒子束管道和顶点探测器突起端20～30厘米的地方，可以额外覆盖那些相对于粒子束以很小角度产生的粒子。接下来是粒子追踪层，称作漂移室（drift chamber），其外形像一层包裹了所有硅层的厚重圆柱形外衣，半径 1.3 米。漂移室通过带电粒子将室中的气体电离进行探测。电离产物会在连接

① 译者注：p 是 positive 的缩写，表示“正”。

在漂移室上的电线所施加电场的影响下向感应电缆漂移，电缆可以将电脉冲转换为电子计数。脉冲时间相对于参照时间（CDF 探测器的参照时间是由加速器时钟提供的）可以得出一个时间间隔，通过该时间间隔可以计算出运动的距离，因此就可以算出位置。漂移室是粒子探测器的一个关键元件，乔治·查派克（Georges Charpak）在 1968 年发明了现代漂移室的前身，他因此获得了 1992 年的诺贝尔奖。在 CDF Ⅱ 的漂移室中，存在 3 万多根感应电缆，是旧探测器中的 5 倍。这种数量的增加减小了电离电荷在遇到感应电缆之前必须漂流的距离，因此可以用这种新的漂移室探测更远的距离。无论旧的还是新的 CDF 漂移室，所记录的粒子位置都能够精确到大约 0.2 毫米。

环绕着硅层和漂移室的是一种大功率的超导螺线管磁铁，它直径为 3 米，长为 5 米。这种磁铁特别大，把一辆家庭轿车放进去后还有空余的地方。该磁铁能够提供一个沿着粒子束运动方向的磁场，以使那些离开碰撞点的带电粒子的轨道发生弯折。根据粒子轨迹的曲率就可以得到它的动量，还可以判断出它所带电荷的符号。

螺线管、漂移室和硅探测器共同组成了定位系统。在定位系统之外，也就是苹果核外的其他部分，是一个近 2 米厚的测量能量的外衣。它由量热器组成，它的工作是测量出现的电子、光子以及粒子喷注的能量。这些电子、光子以及粒子喷注是破坏性的：当粒子几乎毫无阻碍地穿过定位系统后，基本上都会在量热器层停下来，它们的能量在量热器材料中产生粒子雪崩，最后变成一道闪光。

距离碰撞行为最近的是电磁量热器，是由铅和发光材料交替组成的，它将一部分穿越粒子的能量转变成了可探测到的闪光。电子和光子在铅层中会经历韧致辐射和偶产生过程，生成大量的光子、电子和正电子。这些粒子流会在闪烁层中产生闪光，这些

闪光将进入光电倍增管。强子量热器也利用了相似的原理，它是由较厚的钢层和闪光材料层组成的。发生强相互作用的粒子（它们不需要带电）撞击钢材料中的原子核会产生强子流。至少有一些带电粒子可以被闪光物质看到，和电磁元件相比，这种强子流过程的内在不确定性降低了这些元件的精度。

没有几个粒子会通过量热器，这些量热器必须破坏所要测量的每一个粒子。一个明显的例外是 μ 子。它们的质量比较大，这限制了它们通过韧致辐射损失能量，因此它们不会像电子那样通过粒子流来倾泻它们的能量。它们会几乎安然无恙地穿过去，进入设在外层的 μ 子探测漂移室。

将 μ 子挑选出来是很有用的，因为它们是 W 玻色子或 Z 玻色子以及重夸克发生弱力衰变的信号。μ 子计数也能帮助我们避免探测器被数据淹没的问题。改进后的 CDF 探测器每秒能够探测到几百万次粒子碰撞事件。麻烦在于必须把这些测量数据都写入计算机磁带中，以进行存储和等待以后的分析，而磁带每秒钟只能录入大约 50 次碰撞的数据。因此 CDF Ⅱ 探测器和其他粒子探测器一样使用了一个“触发器”，它是某种能够作决定的系统，可以将实验谷粒从谷壳中脱离出来。触发器的职责是决定一个事件看起来是否有意思，如果有意思的话就会启动整台探测器测量并记录它，而其他一些没有意义的事件——数量特别多的垃圾事件——都被简单地忽略了。

进行一次实验，然后忽略实验产生的 99.999%的数据，总是一件不容易处理的事情。CDF 的实验者们所面临的挑战是既要避免丢失有意义的事件，又要避免触发那些无聊的事件。从最简单的层次上讲，假如一个探测器探测到两个 1.5 吉电子伏的 μ 子，它就会触发，探测一个 12 吉电子伏的电子同样也会触发。当探测器探测到大能量流出现，就会采取行动，这种能量流可能是电子或者粒子喷注。通过追踪系统测量到的、在大角度上出现的高

动量轨迹也会符合要求。CDF Ⅱ的一个独一无二的特性，是它具有“b夸克触发”的能力，它能通过硅顶点探测器确定b夸克衰变的证据信号，从而实现这项任务。触发系统总共需要大约25微秒的时间来决定是忽略某个特定的碰撞，还是让探测器集中全部注意力来观测它。

即使在忽略了几乎所有事件之后，CDF Ⅱ每年还是能记录大约3亿次事件，在两年多的运行时间内，它总共产生了大约300000GB（3百万亿字节）的数据，足以塞满上万张普通电脑硬盘。数据以每周600万个事件的速度通过电脑“农场”进行加工，那些程序大部分都是用C++编写的，C++是一种被广泛用于粒子物理研究的计算机语言。

因此这就是发现顶夸克（也包括图4.75中的那个顶夸克）所耗费的代价：大量的实验物理学家，大批技术人员和工程师为其提供支持，一座比一栋房子还要大、重上千吨但建造精度达到微米级别的探测器，大量高速而复杂的电子仪器以及强大到足以能将普通计算机气哭的计算机处理能力，数亿美元的资金，还有一台强大的粒子加速器。就这么多了。

六夸克宇宙

因此，总共存在6种夸克：上夸克、下夸克、奇异夸克、粲夸克、底夸克和顶夸克。把它们都摆到一个表格里面比较它们的质量时，就会显得特别清晰并且更有启发性，见表4.5。完成这件事就不能忽略质量的概念，我们需要给它贴上防止错误的说明，接下来还会在第五章中进一步讨论这个问题。一种说明这些质量的有趣方法是画一些球，使它们体积的比例与各种夸克质量的比例相同，见图4.76。这并不是想说明夸克物理大小之间的差别会以这样简单的方式出现，只是想说明夸克质量范围的差别巨大。这些质量值以及巨大的质量差别怎样来解释呢？这是一个好问题。

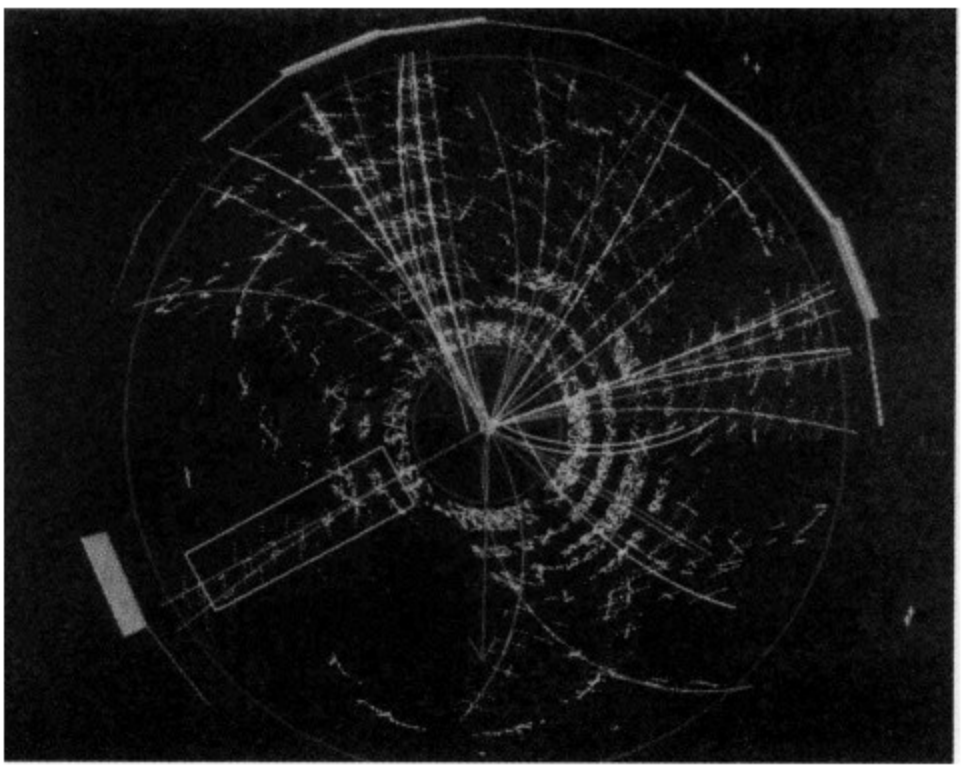
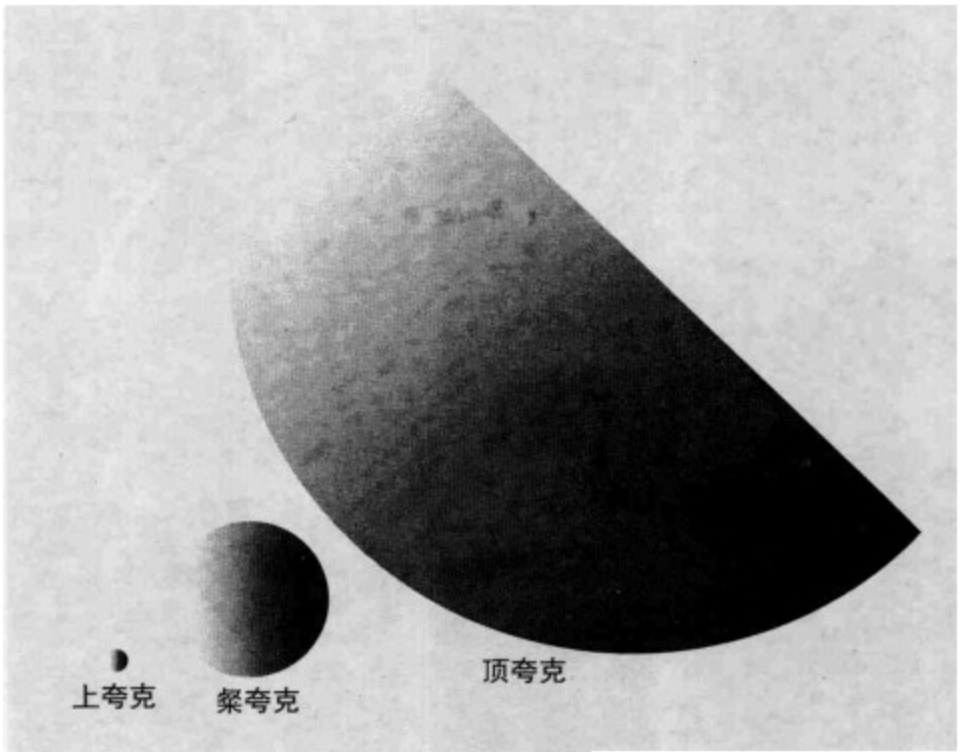
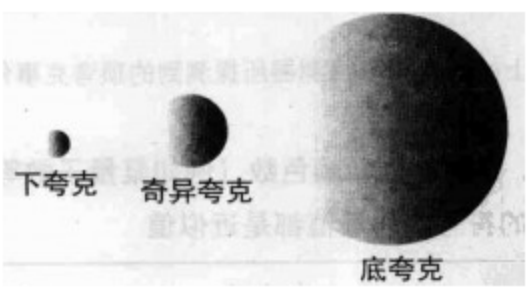


图 4.75　显示在电脑屏幕上的一次 CDF 探测器所探测到的顶夸克事件

表 4.5　6 种夸克的性质。按照惯例，颜色数（例如粲量子数等）与电荷具有相同的符号。质量值都是近似值

夸克性质	夸克味					
	u	d	s	c	b	t
电荷	+2/3	−1/3	−1/3	+2/3	−1/3	+2/3
质量（吉电子伏）	0.004	0.007	0.135	1.3	4.2	174
同位旋	+1/2	−1/2	—	—	—	—
奇异数	—	—	−1	—	—	—
粲量子数	—	—	—	+1	—	—
底量子数	—	—	—	—	−1	—
顶量子数	—	—	—	—	—	+1
重子数	1/3	1/3	1/3	1/3	1/3	1/3





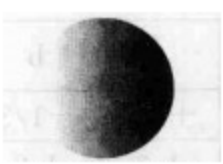


图 4.76　夸克质量。用球体体积之间的差别来表示夸克质量之间的差别，强调3种重夸克的巨大质量。这幅图并不是想说明真实夸克大小的差别会以这种简单的方式出现

另一个好问题是，还存在没有发现的其他夸克吗？有人已经

给它们起好了名字，正等着它们出现：b′夸克和 t′夸克，所带电荷分别为－1/3和＋2/3。随着这些夸克的产生，将会产生与之相对应的轻子，它们共同组成了第四个家庭。最近一段时期，假设一种新的夸克，或者沿着 τ 子的路线假设一种新的轻子，或者假设一种新的中微子，都意味着假设一整个新的家族存在，或者像人们有时称呼它们那样，叫做新的“一代”。但是并没有几个人认为确实应该存在更多的夸克。在过去的时间里，引入粲夸克解决了一个已知的问题——一种弱力衰变的缺失，并且产生了六夸克能够理清 CP 破缺的含义的想法。马丁·佩尔无意中发现了 τ 轻子，因此人们都适时地相信，它的家族的其他成员都是必须存在的。但现在不存在已知的普遍性难题，也不需要增加更多的夸克或者更多的重轻子，或者额外的中微子来解决。因此至少暂时，我们拥有了一个具有 6 种夸克、3 个家族的宇宙，见图 4.77。

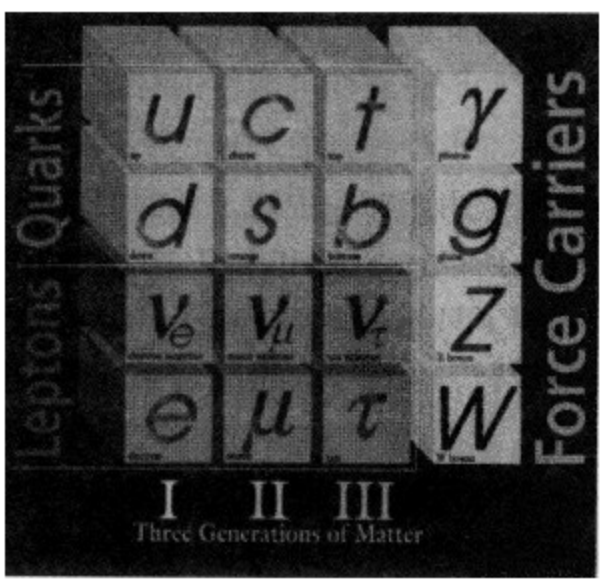


图 4.77　轻子和夸克的 3 个家组，或者称为三代，加上标准模型中携带力的粒子。图片来源：费米实验室

当然，也没有任何现象劝说理论物理学家们不要去假设更多

的夸克。实验物理学家们还在不断地寻找它们。顶夸克刚刚被发现，现在看起来不太可能会有另一种夸克自己暴露出来。但是有一些理论物理学家提出 b′夸克的质量可能会比顶夸克小一些，因此如今的加速器是能够发现它的。当然实验物理学家们已经用 Tevatron 和 LEP 寻找过了，但是什么也没有找到。他们将会用改装过的 Tevatron 以及 LHC（等它再次投入使用之后）再次进行搜寻。如果确实存在另一夸克的话，具有很高精度的 B 粒子工厂说不准也能发现一些蛛丝马迹。

是否存在相反的论证，能够通过某种实验来限制夸克家庭的数量呢？答案是肯定的，人们已经在做这件事了，尤其是在 SLAC 利用线性加速器以及在 LEP 上进行的利用电子-正电子碰撞实验研究 Z 粒子产生的过程。Z 玻色子可以通过许多方式衰变，例如衰变为像电子和正电子、μ 子和反 μ 子这样的轻子，但它更容易衰变为夸克，因此就会产生强子。将各种测量到的衰变可能加起来，然后将它们同总衰变可能性进行比较，结果发现差了很多。这种差额同标准模型预言的 Z 玻色子只能衰变为 3 种类型的中微子的情况相符，由于存在中微子观测的难题，这些中微子衰变率是看不到的。所以在 SLC 和 LEP 中积累出的上百万 Z 玻色子衰变事件，说明了只能有 3 种中微子。或者也可以说，实验结果指出存在 3 种“轻”中微子，也允许第 4 种中微子存在，但它的质量要超过 Z 粒子质量的一半，这样它就不会在 Z 粒子衰变过程中表现出来。因此，如果不存在这么重的中微子——但一般来说，人们在研究假设中还是认为所有的中微子都有极小的质量——按照家族模式，中微子数量对应着夸克对的数量，那么打赌说只存在 6 种夸克和 3 个家族是很保险的。

简单地说，夸克种类的数量不会对 QCD 理论造成很大的影响，将它们结合在一起的强力同它们的味道无关。而家族问题或者叫代的问题，也就是夸克味道的数量，在很大程度上是标准模

型中有关弱相互作用部分的内容，也是那些超出标准模型理论要讨论的问题。

QCD 理论可能是“味”盲，但它也并不是完全与味无关。尽管在最简单的意义上，只存在一种夸克味也可以满足 QCD 理论，但这会导致非常奇怪的强相互作用世界。例如，大多数介子将会在只有一种味的宇宙中没有立足之地，重子也是这样，这两类粒子都只能被压缩到基于一种夸克的激发态上。根据 QCD 理论得到的许多结论，具有不同质量的不同夸克以及许多试图解释介子和重子结构的努力都变成了多余的。我们今天所了解的质子将不复存在，取而代之的是一些仅仅由一种味的夸克所组成的重子，有点像 Δ^-、Δ^{++} 以及 Ω^- 等。中子将不会存在，因此大概也就不会存在用强力来解释的原子核了。

当理论物理学家们说明 Tevatron 上一个质子和一个反质子对撞时，他们会将质子和反质子都描绘成本质上不发生相互作用的夸克包，但这些夸克包在某种程度上依赖于存在多少种味道的夸克。此外，联系实验结果和理论计算的正是夸克的味数，因为反应是通过碰撞粒子中不同的夸克和反夸克进行的。几乎每个要和基于 QCD 理论的某些预言相比较的实验都是这样。

但是在 QCD 理论和夸克味道数之间存在着更加基本的联系。一个胶子能够在很短暂的时间内舒展一下它的量子“身躯”，变成一个夸克－反夸克对。更多的夸克味给胶子提供了更多的选择，夸克的味数正是通过这种方式，同 QCD 材料以一种不可忽视的方式交织在一起。根据 QCD 理论，强相互作用的强度实际上就依赖于夸克味道的数量。甚至连决定 QCD 理论影响范围的基本长度尺度也与味数有关——QCD 理论可能无法识别味道，但是味道数的确是 QCD 理论的一个基本数字。

第五章 QCD 理论的一个数字

传统的藏茶是一种需要逐渐习惯的佳茗。浓郁而醇厚的茶香会萦绕许久，然后慢慢散去。喝的时候，要先将茶放在包着黄铜的木制搅拌器中，加入咸黄油后搅拌，然后重新加热，才可以开始品尝。

黄油是用牦牛的奶制成的，它会带来某些意想不到的味道。但是味道中最为重要的一个因素就是水，尽管水已经烧开了，但这种水温，还不足以冲开英国绅士们都很熟悉的、在布满玫瑰花瓣的小别墅中享用的那种茶叶。关键问题在于，西藏位于高原之上，其高度足以使当地的气压比海平面处的气压低得多。因此水会在很低的温度下沸腾，其温度低到能让沸水中的小虫子都安然无恙，但肠胃不好的游客们却遭了殃。

尽管水的沸点随着气压的降低而降低，但也会随着气压的升高而升高。换句话说，水从液态水相转变为水蒸气相（这称作相变）的温度，随着气压的升高而升高。但这种简单的关系不会无限制地继续。在 218 倍于标准大气压的气压下，水会在 374℃沸腾。如果气压再升高，水就不存在沸点了：气相和液相变得无法区分。气体和液体之间差别消失的那一点，其温度和压强就称作水的临界点。

另一个临界点的例子，也是凝聚态物理学家们比较喜欢的一个例子，就是铁磁性。在很高的温度下，铁并不存在固有的磁化倾向。但如果铁被冷却，到达770℃以下，就会出现磁性。在该温度以下，不同区域（称作domain，磁畴）中，每一个单独的原子磁体会自发地排列起来，它们共同的有序力量最终战胜了无序的热运动。一般来讲，同一个样品中的不同磁畴具有不同的磁化方向，但是在室温下，受磁场影响的铁块中，至少有一些磁畴的磁化方向会随之发生改变，或者某些磁畴的范围会通过吞噬别的磁畴而扩大，因此铁块会具有很强的磁性。在临界温度770℃处，那些人们熟悉的室温下的铁块，会从有序的“铁磁”相变成无序的“顺磁”相，这在许多金属（例如铝）中都很常见。还存在其他许多这种所谓临界现象的例子，在这些现象中，物质两相之间的差别将会消失。

就像肯尼斯·威尔逊（Kenneth Wilson，图5.1）本人曾经提到的那样，两个相邻的水分子不知道它们是在一个茶壶中，还是在太平洋的中心。在单个原子行为占主导地位的尺度下发生的过程，对在比它大几十亿倍的茶壶或者海洋的尺度上发生的行为，几乎没有任何影响。但是有一种情况下，分子与其临近的分子之间的相互作用，确实会影响整个茶壶的组成，从最小的尺度到最大的尺度都会参与其中，这就是临界点的情况。

对于一个盛水的容器，其内部的压强为218个标准大气压，随着温度逐渐升高并且接近374℃，“水”会变成由各种大小的水滴和气泡组成的混合物。在临界点，这些气泡和水滴互相纠缠到一起，其大小从分子尺度到容器尺度都有。它们在整个容器内形成了一个完整的密度变化谱。每一个分子都能受到所有其他分子的影响。

在这样一种长度尺度下，各部分对整体的贡献是一个很难处理的问题。通常，小的起伏会通过某些方式达到平均，物理学家

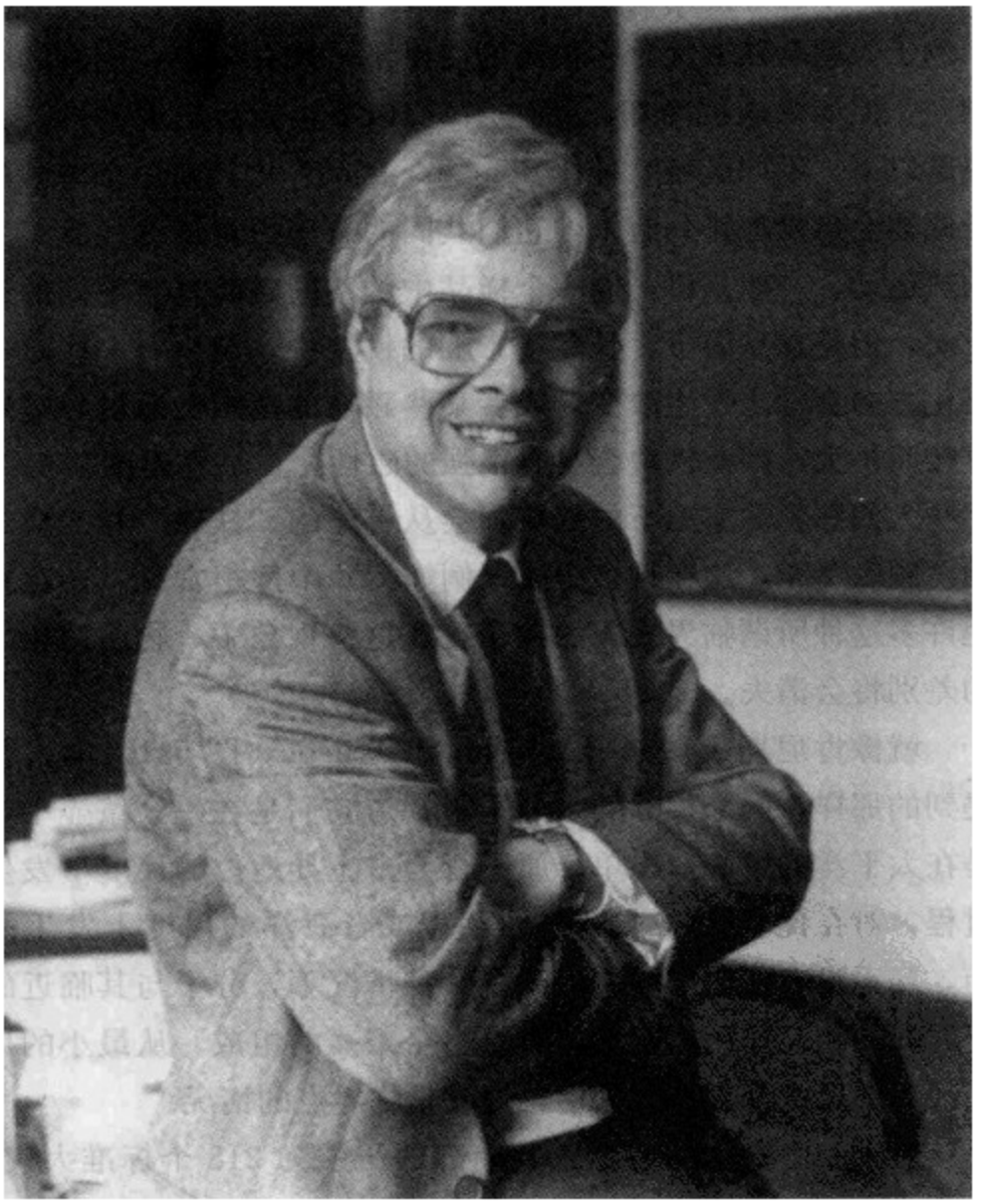


图 5.1　肯尼斯·威尔逊于 1998 年摄于俄亥俄州立大学的办公室中。威尔逊与重正化群相关的思想，给 QCD 理论以及其他方面带来了深远的影响。他因在“与相变有关的临界现象理论”方面的工作获得了 1982 年的诺贝尔奖。图片来源：肯尼斯·威尔逊

们很早就知道怎样处理这样的问题。但是对于那些所有长度尺度

都参与贡献的现象，用平均的方法会得到错误的答案，处理它们需要一些新的方法。这就是重正化群方法。

将重正化群用于临界现象研究，它会联系着系统不同的“放大率”，就好像将一台强大的摄像机推得越来越远，以能够将越来越大的景物都包括在镜头当中，而比较小的细节将会逐渐从视线中消失。重正化群描绘了从纹理精细的图像，连续地演变到纹理粗大的图像的过程——它联系了不同尺度上的物理。在临界点上，像密度起伏或者磁场一致区域这些属性，随着摄像机的远近移动，呈现出所有的长度尺度，而镜头中场景的整体外观没有任何改变。换句话说，对于在临界点上，通过某种流体中的液滴或气泡的大小，或者是在临界温度下，通过一块铁中那些具有相似磁化方向的区域所形成的簇的性质，不可能断定其放大率的设置是高还是低——两者看起来都一样，至少在统计上是这样。事实上，如果只通过目镜进行观察的话，你不可能知道显微镜的放大倍数设定在哪个值。这是一种不变性，因为在距离发生了变化后，一切都不发生任何变化。这种尺度不变性是一种对称性，就像在重正化群中通过利用“群”这个词所表达的一样。

将重正化群应用于临界现象，允许理论物理学家们预言一些可以测量的性质，这些性质依赖于系统的所有长度尺度。例如，对于水来讲，如果在外部对其施加压强，它的体积就会发生相应的变化；这种可压缩性在临界点会变得无穷大，在接近临界点的时候，压强与体积会按照某种特殊的数学关系进行变化。对于一块简单的铁磁性磁铁，材料的磁化率随着外加磁场的变化也满足同样的数学关系，在临界点也同样会变成无穷大。这两个例子说明了这两种临界现象的主要性质。首先，在临界点附近，它们表现出对描述系统的相应物理量典型的依赖关系，而在临界点上物理参数会出现无穷大的峰值。第二，表面上差别很大的物理系统，在临界点都有一种普遍的特点，在上面的例子中，液-气转

化和铁块中的磁化现象表现出同样的变化方式。

人们通过用重正化群研究临界现象得到的这些概念，促成了现代热物理研究中的一场伟大胜利，其中包括流体和磁性系统在临界点的行为、转变为超导态的临界点行为、油和水的整体混合物在高温下临界点的行为和其他很多现象。但是从表面上来看，所有这些现象好像都与强力没有太大的关系。但确实存在一种联系，这种联系就是重正化群以及在长度尺度变化下的不变性。

电磁作用变得更强……

家庭大百科全书给出的电子电荷值，大约是 1.602177×10^{-19} 库之类的数字。尽管这是一个十分正确的数字，但从某种程度上来讲，它有点过于简化了。发生相互作用的粒子的电荷和质量，并不是能刻在石头上的那种一成不变的值。这些量的值实际上要依赖于观察者是谁，或者更准确地讲是依赖于他们看得有多仔细。对于电子和 QED 理论来讲，这只是一件新奇的事物。但对于 QCD 理论，这却是最基本的概念。

戴维·科尔提克（David Koltick）和他在 TOPAZ 小组的同事，利用现在已经停用的 TRISTAN 电子-正电子对撞机［位于东京附近的日本高能加速器研究机构（KEK）］，通过在高能条件下观测一个电子的有效电荷，从而检验 QED 的奇异角（curiosity angle）。这可以通过对比两个过程 $e^+ + e^- \rightarrow \mu^+ + \mu^-$ 和 $e^+ + e^- \rightarrow e^+ + e^- + \mu^+ + \mu^-$ 来完成，见图 5.2。第一个过程只涉及一个光子，并且通过光子传递的动量也很高，因为电子和正电子结合的全部动量都是通过它流出去的。所以在这种情况下，电子和正电子之间的相互作用是短程的。第二个过程包含两个光子，与第一个例子相反，通过这两个光子传递的动量都很小。做一个不太恰当的类比：可以想象一下两条小河在某处汇聚成一条大河的情

况，再想象一下连接这两条大河并且与大河并排流动的小溪的情况，然后将这两种情况中的水流做一下比较，就类似于以上两种情形。同第一个过程——电子-正电子湮灭过程相比，第二种相互作用的距离相对长一些。

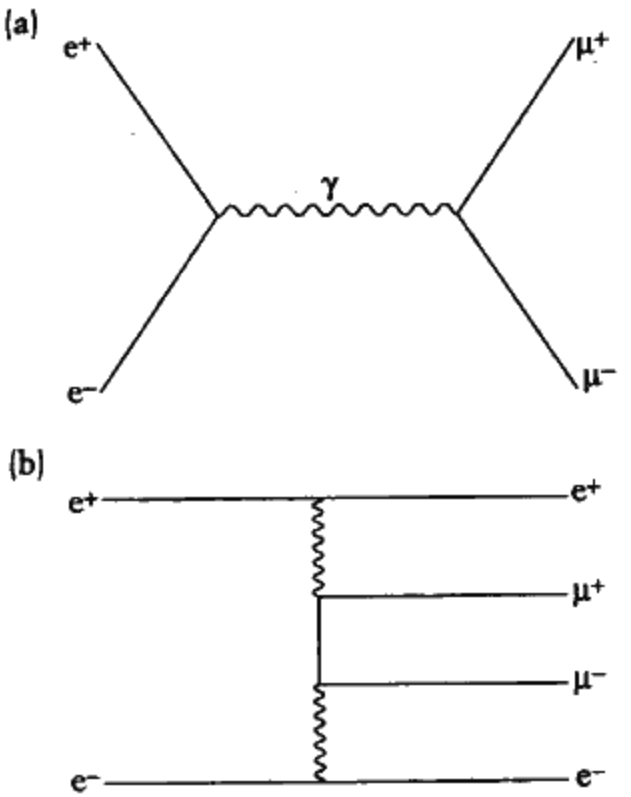


图5.2 过程 $e^+ + e^- \to \mu^+ + \mu^-$ [图 (a)] 和过程 $e^+ + e^- \to e^+ + e^- + \mu^+ + \mu^-$ [图 (b)] 的主费曼图

通过比较这两个过程的两种散射率，TOPAZ 小组得到了一个在高能情况下的电磁耦合强度值 α。而另一方面，凝聚态物理有一些能在低能情况下确定 α 值的精确技术——忽略一般的误差，现在人们普遍接受的 α 值是大约 1/137.036——很可能由于历史原因，人们经常将 α 值表示为 1/137。TOPAZ 小组在他们于 1997 年发表的论文中报告称，他们在迁移动量为 57.77 吉电子伏平方（GeV^2）的条件下测得的 α 值为 1/128.5，这个值在误差允

许的范围内，同理论预言值 1/129.6 相符。

TOPAZ 小组所看到的是两个带电粒子，它们之间的距离越近，它们之间的电磁耦合强度就会越强。他们的实验绝不是第一次观测到这种效应的实验，但这次观测却十分清楚和明确。可以通过这样一种物理图像来考虑问题：假设有一裸露的电荷，周围环绕了一团虚光子云——这同用来修正库仑定律的物理图像是一样的，库仑定律的修正可以从氢光谱中的兰姆位移现象体现出来。在任何一个时刻，这些虚光子中的一部分可能会变成粒子-反粒子对，例如电子和正电子对，一般带相反电荷的虚粒子会离裸露的电荷近一些，见图 5.3。其实际效果就是裸电荷被一团虚粒子屏蔽起来，这就是被屏蔽的电荷在低能条件下的情形。但是如果用一个高能探针击打电子，也就是说，将探针插入屏蔽层越深，从而使裸电荷暴露得越多，此时电磁力强度看上去就会增加。

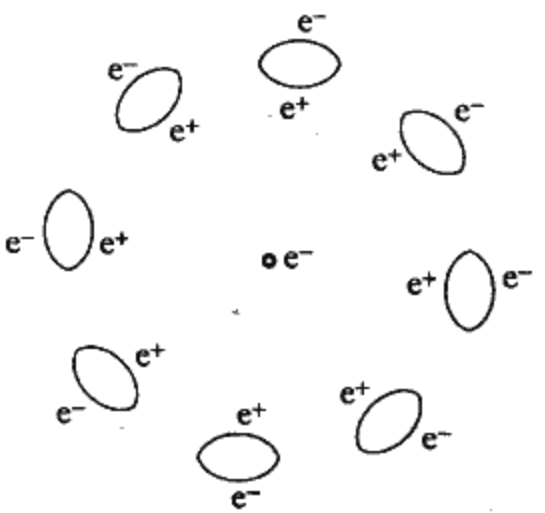


图 5.3　一个裸电荷——位于中心处的电子怎样受到真空中产生的电子-正电子对屏蔽的图示，图中用圆圈表示电子-正电子对

可以想象一个专制的独裁者，他的思想特别偏激，周围的议员都对残暴的统治哲学感到反感，他们会执行他的命令，但会稍

微缓和一下他的计划。执行独裁者意图的那些政府文员阶层，会将其进一步地缓和，因此公众所看到的是这个独裁者被冲淡后的形象。就像议员和文员屏蔽了那个“裸露的”、令人憎恶的独裁者，这样公众就会看到一个较温和的形象一样，虚粒子对屏蔽了一个裸露的电荷，这样从外面看上去，就会觉得电子所带的电荷相对弱一些。如果刺穿了以上任何一种屏蔽层，裸露的核心就会暴露。

TOPAZ实验中的屏蔽效应，与重正化过程中屏蔽效应的机制是一样的。重正化确定了一个不可测量的裸粒子与一个有限的、可测量电荷与质量的粒子之间的关系。严格地讲，为了找到这一关系，重正化方法引入了一个重正化标度，在“短”距离和“长”距离之间任意确定了一条分界线。但物理不会关心这条分界线的确切位置，因此理论物理学家们预言，长短距离的分离点所在的位置，肯定不会随标度的变化而改变——这是一种尺度变化不变性。那么这种重正化标度不变性到哪里去了呢？它被理论中一些关键的物理量吸收了，比如粒子的电荷或质量等。因此这些量一定会随着长短距离之间的分界线（也就是重正化标度）的变化而变化。通过研究物理上这种长度尺度变化下的不变性，重正化能够明确说明那些关键的物理量是怎样变化的并在不同的重正化标度上确立了电荷与质量之间的精确关系，因此也就确立了它们在不同能量下的精确关系。在这里，重正化说明，一个电子的电荷在另一个入射的带电粒子看来有多大，要依赖于这两个粒子之间的相互作用有多密切。这种耦合强度随着能量演化的事实，也说明了“耦合常数”这种表达方式是一种不恰当的说法，因为它根本不是一个常数——物理学家们已经发明了另外的名词来替代它，像“有效耦合常数”或者更确切的“跑动耦合常数”，以表示这种能量依赖性。

尽管重正化允许理论物理学家们在不同的能量下计算耦合常

数，但是我们还忽略了一些东西——实验不得不提供一个电荷初值，比如可能是在低能情况下的测量值。这有点像，已知离开一个拥挤的房间的人的频率是多少，而想要知道现在房间里还剩下多少人，但是却不知道一开始房间里有多少人一样。要得到答案，就需要给出一条额外的信息——一个绝对值来确定整体的标度。这样，对于人们离开房间的问题，只要知道了最一开始时刻房间中的人数，留在房间里的绝对人数也就很容易算出来了。电子电荷的问题也是一样，一旦通过测量，在低能状态下确定了电荷值的大小（例如用凝聚态物理技术确定的电荷值），重正化群就会提供一种计算高能条件下跑动耦合常数的方法。这就是 TOPAZ 小组经常用到的方法，他们用这种方法计算出耦合常数的理论计算值，然后同他们得到的实验值进行比较。

还有一点也很明显，那就是在 QED 理论中还有一些没有完成的工作。耦合常数会不断增加，这种增加导致了无限增大的迁移动量，这说明会出现一个点，耦合常数在这一点上并不是很小，因此微扰论在该点会没有意义，因为微扰论希望得出一个逐项递减的幂级数结果。换句话说，像在第三章已经提到过的那样，微扰论最终在 QED 理论下也失效了。但这是在能量特别高的情况下发生的——大约为 10^{26} 吉电子伏——这完全超过了所有粒子加速器的能量范围。实际上，在如此高的能量之下，传统的时空观念本身也没有意义了。因此，这种在 QED 范围内的失效实际上只存在于理论范畴。

……而强力却变得更弱

所有有关 QCD 理论的研讨会，如果没有提及强相互作用耦合参数 α_s 的最新测量值，都是不完整的。

物理学家们已经研究了许多粒子反应，想通过这些反应来寻

找一个可靠而精确的 α_s 值，其中最受他们欢迎的一个反应仍是 τ 子的衰变过程。这些 τ 子能够衰变成大量的中微子、反中微子、μ 子和电子。τ 子本身是通过电子-正电子湮灭产生的，见图 5.4。但是 τ 子也很重，其质量足以能通过 W 玻色子衰变，成为中微子

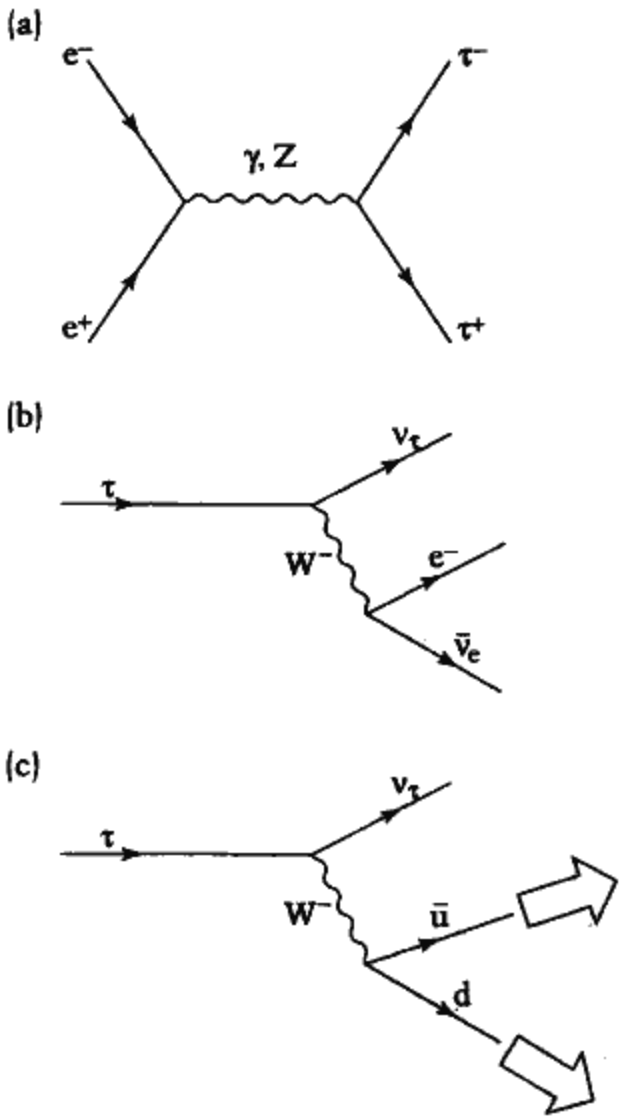


图 5.4 τ 子在电子-正电子碰撞过程中产生［图 (a)］。τ 子能够通过其他轻子进行衰变［图 (b)］，或者衰变为夸克-反夸克对，从而形成强子［图 (c)］。通过在图 (c) 的基础上再包括更高次的费曼图，将轻子衰变和夸克衰变的衰变路径进行对比，可以得到强耦合参数的值

和发生强相互作用的粒子，见图 5.4 中的图（c）。最初看上去，这可能会有点儿令人吃惊，但实际情况是这样的，夸克是自然界中唯一的一种能感觉到自然界中所有作用力的粒子，它特别喜欢同 W 玻色子发生弱相互作用。理论物理学家们之所以喜欢 τ 子，是因为它是一个测量强耦合参数的好帮手，因为他们能将 τ 子计算到包含 3 个圈图的费曼图。这大致相当于在 QCD 理论中可以达到的最精细的水平了。理论物理学家们喜欢它的另一个理由是，他们不能确定怎样计算的那些部分都会集中在一起，可以将它们称为“非微扰效应”，它们的值特别小，不会对实验的基本结果造成影响。

许多实验小组，例如 CESR 电子-正电子对撞机的 CLEO 合作小组，已经公布了基于 τ 子衰变的强耦合参数测量值。其结果是 $\alpha_s(m_\tau)=0.35\pm0.03$——$\alpha_s$ 没有单位，因为它仅仅是一个数。下角标 s 代表“强力（strong）”，以将 α_s 同 QED 理论中的 α 区别开。与 QED 理论中的 α 相比，α_s 存在着巨大的误差。上述 α_s 值的结果中，还有另一个引人注意的特征，就是有关（m_τ）的说明。这是因为强耦合参数强烈地依赖于能量。因此说明能量尺度是必要的，在上面例子中的能量值是 τ 子的质量，1.777 吉电子伏。

如果将从 τ 子衰变得到的 α_s 结果，同其他实验得到的结果加以比较，就能够很明显地看出（m_τ）标签存在的必要性。这种测量值太多了，以至于无法在这里全部列出来。还有另一种测量 α_s 的好方法，那就是研究能产生不同粒子流（称作喷注）的碰撞。例如 HERA 的 ZEUS 合作小组和 H1 合作小组，都将观测电子-正电子碰撞所产生的粒子喷注作为一种测量强耦合参数结果的方法，见图 5.5。其中一种粒子喷注是质子被撞碎后的剩余物，另外两种分别对应于夸克和胶子［图 5.5 的图（c）］或者夸克和反夸克［图 5.5 的图（b）］。它们得到的复合结果是 $\alpha_s(m_z)=0.1179\pm$

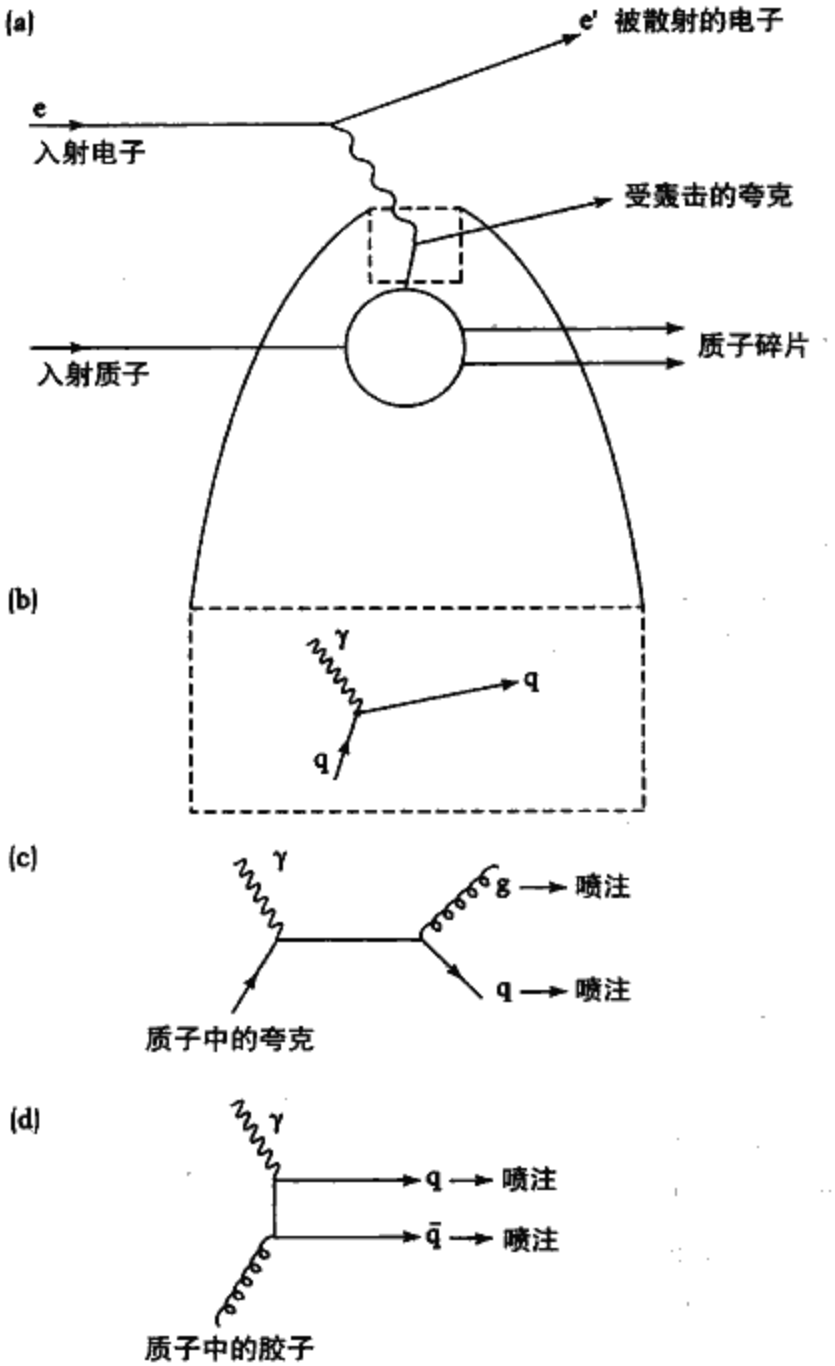


图5.5 电子-质子散射过程中的三喷注事件。图（a）表示普通的深度非弹性散射。子过程 $\gamma q \to q$ 用虚线框表示，并且在图（b）中被详细加以说明。对于三喷注事件，该子过程被图（c）和图（d）中所示的过程所代替，这两个过程除了能产生质子碎片喷注之外，都会产生另外两个喷注

0.006。这个数字看起来确实与 0.35 存在很大的差别，它反映了强相互作用耦合对能量的依赖关系。因此在引用 α_s 值的时候，包含能量信息是很重要的。

实验物理学家们已经积累了特别多的 α_s 测量值。当把这些测量值加以比较，就会发现一些性质。其中一种性质就是，用各种不同的方法都会得到一致的结果——如果不是这样的话，QCD 理论肯定就会存在问题。另一个特征就是，强相互作用的强度并不依赖于夸克的种类，由于强力与色有关而与味无关，因此这也是一个预料之中的结果。最后，随着能量升高，强相互作用的耦合强度确实像渐近自由理论预言的那样减小。

考虑到能量与耦合强度之间这种确定的关系，在引用强相互作用耦合常数的测量值时，人们经常用在某个特定能量尺度下的对应值来表示，习惯上经常将能量取为 z 玻色子的质量 m_z。这样，实验物理学家们就可以将不同实验得到的强耦合常数测量值进行比较。例如由 τ 子衰变实验得到的 $\alpha_s(m_\tau)=0.35\pm0.03$，在更高的能量下它会演变为 $\alpha_s(m_z)=0.121\pm0.003$，这同上面 HERA 得到的结果一致。对许多的实验结果进行平均，就可以得到强相互作用耦合常数的值，其中最好的一个选择就是取 $\alpha_s(m_z)=0.117\pm0.002$。

实验物理学家们证明了，在 QED 理论和 QCD 理论中，相互作用强度都依赖于能量。它们尽管都是依赖于能量的相似系统，但对于能量变化的反应方式却相反。在 QED 理论中，屏蔽提供了一个简单的物理图像，表现了随着被测量的粒子同探测粒子之间的距离越近，它们之间的耦合强度就会越强。与之相反，在 QCD 理论中耦合强度却会降低。这种现象可以用反屏蔽效应来解释。一个裸露色荷周围的虚粒子对，实际上会使夸克上的色荷显得比它本身更强，当一个探测夸克刺入反屏蔽层，其核心部分就会显得更弱，而不是更强。一个电子会在它的周围产生环绕着它

的相反电荷；与之相反，一个夸克看起来会在自己周围产生与它的色荷类型相同的色荷。这种差别的根源在于携带强力的粒子——胶子本身也带有色荷。因此，在裸夸克周围的虚粒子云中，出现的胶子会有直接贡献，而 QED 理论中的光子却不会。正是这些胶子的贡献比夸克-反夸克对的贡献还大，并且与夸克-反夸克对的作用相反，它们最终取得了胜利。如果自然界为我们提供了更多种类的轻夸克，那么夸克的贡献就会胜出，宇宙将会与现在的样子大不相同。

这种反屏蔽效应不存在简单的物理类比，但是前面所举的独裁者的例子可能会给我们一些帮助。想象一个仁慈的统治者，像上面那个例子一样，他周围也有一些议员，而这些议员周围又有一些政府官员。但这一次，各个部门之间的交流都是通过恶毒的通信者——胶子来完成的。他们通过自己的日程、秘密组织以及联系方法来施加他们自己的影响，控制了议员及政府官员对民众的影响，在外界看来，他们的统治者是一个彻头彻尾的暴君。但是在接近他本人后，你就会发现，他实际上十分温和。

在 QCD 理论中，强耦合参数 α_s 与能量之间的联系，可以用重正化群辅以微扰论方法来确定。α_s 最终的表达方式是整个 QCD 理论中最关键的结果之一。乍一看上去，该结果同 QED 理论中对应的方程很相似。但再仔细一看就会发现，分母中有一部分看起来很奇怪。这就是 $33-2n_f$ 那一部分，其中 n_f 表示夸克的味数。可能存在的夸克种数竟然出现在该方程中，这是令人吃惊的。但是如果味的数量是 17 或者更多，那么这个因子就变成负的了，那样的话，理论中就不会出现关键的渐近自由特征。这就说明了，为什么如果大自然构建在更多轻夸克之上，宇宙将会变得面目全非。其中的数字 33，与胶子以及色荷群是 SU（3）群有关——不同的对称群将会导致不同的数字。$2n_f$ 源自于夸克，反映了在相互作用过程中，胶子能转变为不同虚夸克-反夸克对数

量的可能性。因此，具有渐近自由性质的 QCD 理论必须有 16 个或更少的夸克类型，虽然这比通过其他方法所得到的夸克数限制要大得多，但这仍是一个对夸克种类的有效限制。

在强耦合系数与能量之间的关系中，还有另一微妙之处。理论虽然说明了耦合强度怎样随能量尺度变化，但它并没有设定尺度的零点。这个问题在 QED 理论中也会出现，但在 QED 理论中的解决方法很简单——在“零能量”下测量耦合常数，然后把所有其他的值同这个值关联起来。换句话说，就是利用简捷实用而又非常精确的技术，来确定一个孤立电子的电荷。

但是在 QCD 理论中，问题就出现了。我们无法像实验者测量一个孤立电子上的电荷那样，来测量一个孤立夸克上的色荷强度。耦合常数在低能的时候确实会变得很大，但低能情况下的测量值，甚至都无法与用微扰论推导出来的结果进行比较，因为微扰论方法只在高能情况下有效。物理学家们所能做的只是在某些合理的能量下测量 α_s 值，从而确定整体关系，就像在 QED 理论中，知道一个电子所带的电荷满足与电磁耦合常数正向递增的关系一样。

实际上，在 QCD 理论中，人们更习惯于用某个参数来表示强耦合常数与能量之间的关系，只要实验物理学家们测量一下 α_s，这个参数的值就知道了。这个参数，也可以等价地说是在某个特定能量下测量到的强耦合常数 α_s，是 QCD 理论的特征数字，是该理论的基本常数。它必须通过实验来确定——而其他所有重要的量都可以从理论得出。

习惯上将这个未知的参数称作 Λ——希腊字母“兰姆达”，它具有质量的量纲。有人可能会想，QCD 理论的“基本常数”应该有一个自己的、恰当而响亮的名字，但这个参数却注定默默无闻：粒子物理学家们在提到它的时候只是用 Λ、“Λ 参数”或者“QCD 尺度参数”来表示。

从理论上来讲，确定这个基本常数 Λ 是很容易的。例如，在上面提到的 H1 合作小组和 ZEUS 合作小组用电子-质子喷注测量强耦合参数的例子中，两个小组都利用他们的散射数据，同 QCD 理论对该过程的预测值作比较，不断改变 Λ 值，直到理论值与实验数据符合得最好。反过来又可以用它来确定强耦合参数 α_s。

但在实际中，找出一个 Λ 的确定值是很难的事情。主要原因是在比较高的能量下，要有效地降低耦合强度，需要耗费越来越多的能量。换句话说，要想深入围绕着色荷源头的反屏蔽层会变得越来越难。这就表示：要想精确测量 Λ 的值，就需要达到很大的能量范围。在低能情况下，这种望远镜效应不会造成什么问题，但是这时，实验也进入了一个微扰 QCD 理论不再可靠的能量范围。公认的 Λ 值大约是 200 兆电子伏。

相对于一些作为标准点的常数，Λ 参数隐藏得很深。QCD 理论的主方程根本没有 Λ 参数。在一般高能近似中都忽略了夸克的质量，这被人们普遍接受——QCD 理论的主方程中没有提到任何有关质量、长度或者固有大小的东西。QCD 理论的这种（与质量无关的）主方程根本不了解世界的大小，而人们却想要把它应用到这个世界中。理论正是通过这种简单的方式实现了尺度不变性——也可以把这种理论应用于行星那么大的粒子，这当然与实际情况有着天壤之别。

但是当在 QCD 理论的主方程中同时进行量子化和重正化时，就会发生一些令人吃惊的事情。重正化引入了一个长度尺度——动量截断，或者等价地说是一个重正化尺度。实际上，正是理论中物理内容关于这种尺度因子的不变性，导致了强耦合参数随能量的改变。但是要想得到强耦合参数的实际值，就要打着 Λ 参数的幌子引入重正化尺度。QCD 的量子理论因此也以 Λ 参数的形式依赖于尺度：QCD 主方程的尺度不变性，在理论转变为一个完整量子理论的过程中做出了牺牲。200 兆电子伏的 Λ 值对应着一

个大约 1 飞米的长度尺度，这比 π 子的波长稍微小一点，大致相当于质子的大小。换句话说，Λ 值将理论家族推入了亚原子核领域。QCD 的量子理论是一套不能应用到行星或者台球上的理论。

如果能量尺度接近 Λ 值，那么强耦合参数就会很大，微扰论就不再适用。换句话说，Λ 值能够表示微扰理论将在什么时候失效。这是耦合参数-能量图的低能端，见图 5.6。在高能情况下，耦合参数很小，微扰论是有效的，因此各种有效的微扰理论计算都可以进行，例如那些有关深度非弹性散射过程的计算。这就是微扰 QCD 理论能够应用的范围。大多数 QCD 理论都是微扰 QCD 理论。在大约 1 吉电子伏以下的低能世界中，是 QCD 理论

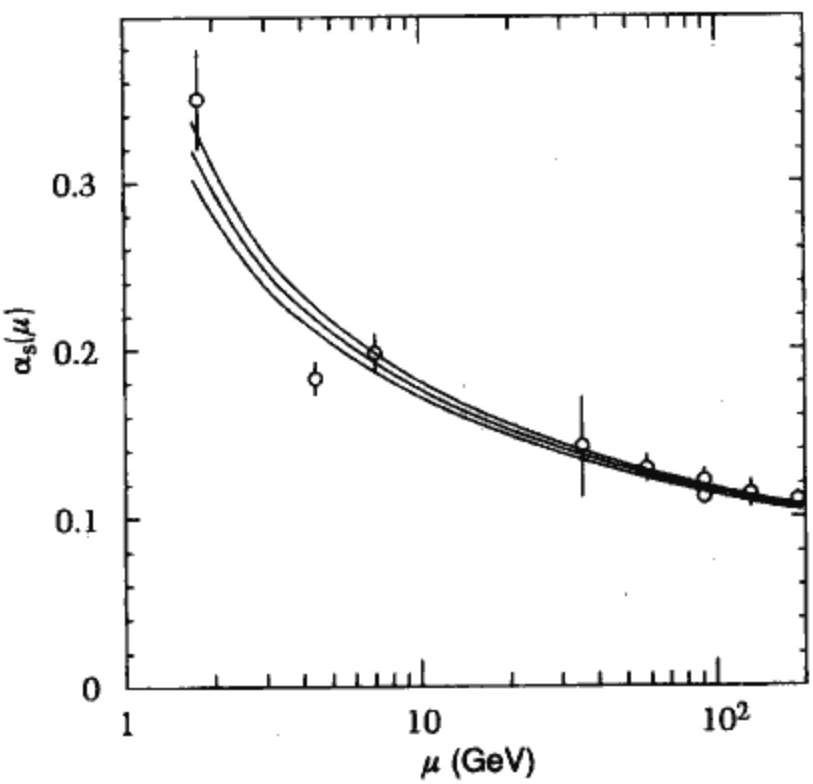


图 5.6　随着能量增加，强耦合参数 α_s 的强度会降低，这是一种与众不同的现象。上面和下面的曲线说明了中间那条线的统计分布范围。图片来源：粒子数据组（Particle Data Group）

隐藏的山谷，是一块不易探索并且越来越难走的区域。但是有一群特殊的人，在这些荆棘丛中成功地开辟出了几条道路。他们就是所谓的“格点 QCD”的研究者。他们的目的就是直接利用大量计算机，用数值模拟的方法解 QCD 方程，从而直接解决这种 QCD 理论“非微扰”区域的问题。格点 QCD 是第十章要讨论的主要问题。

在非微扰 QCD 理论茫茫黑暗中的某个地方存在着禁闭，这是高能情况下渐近自由的对立面。能量很低的情况下耦合强度就会很大，这就说明了为什么没有人能看到自由夸克和胶子。它们结合得太紧密了，都被禁闭在很小的一个区域之内。

一个高深莫测的数字

作为一种基本的参数，Λ 显得极为讨厌。一些人可能希望找到一个简单的、能够明确定义并很容易测量的数字，但是他们没有那么幸运。要从实验中得到 Λ 值是一件难得不能再难的事情。

一开始，Λ 要受到前面说的能量望远镜效应的影响，但这仅仅是一个冗长而不幸的故事的开始。一般所用到的计算，最多只包含一个或两个圈的费曼图。但是在这种复杂程度下，联系强耦合参数和 Λ 的关系不是唯一的。实验物理学家们希望推导出一个 Λ 值，它一定要能详细地解释他们的强耦合参数与 Λ 值是怎样联系在一起的。

在实际计算中，不可能包括每一个能想到的复杂情况以及所有可能的费曼图，这又产生了与之相关的一个问题。可怜的物理学家们不得不在某个地方停下来，通常是包含两个圈的费曼图，或者偶尔有一系列包含三个圈的费曼图。问题在于，尽管一个物理上可测量的量（比如说散射截面）不会依赖于理论物理学家们所选择的重正化方法，但这对实际计算却会产生影响。因此，尽

管在用两种不同方法计算一串无穷多的量时，最终能得到相同的结果，但是如果算到一半停下来，然后比较两种方法得到的结果，就会发现重正化方法的选择会引起差别，这就是方法选择对计算细节所带来的影响。为了说明这种思想，我们可以将数字 1 用两种方法拆开，一种是 $1=1/2+1/4+1/8+1/16+\cdots$，另一种是 $1=1/2+1/6+1/12+1/20+\cdots$，它们分别使用了级数 $\frac{1}{2^n}$ 和 $\frac{1}{n(n+1)}$，其中 n 是比零大的正整数。如果计算所有项的话，右边的两个分数序列加起来都等于 1，并且两个级数的第一项是相同的。但是如果将两个级数都截断，比如说都截断到第四项或者是第五项，或者更多项，然后把它们加起来，这两串数将会得到不同的答案。对于 QCD 理论，这种截断就意味着对于所有最简单的计算，两个理论物理学家预测某个特定过程时，所写出的方程也可能是不相同的，这依赖于他们每个人所用到的重正化方法。这就是为什么理论物理学家们要约定一个惯例，以在进行理论计算值之间以及计算值与实验之间的比较时减少复杂程度的原因之一。

最后，还存在夸克味数的问题。发生相互作用的能量说明，不同的夸克味在该过程中起了重要作用。在该过程中，大量的虚粒子会轻而易举地产生和消亡。上夸克、下夸克以及奇异夸克都特别轻，因此很容易产生。粲夸克在某些能量比较高的相互作用中也很容易产生，在足够高的能量下，低夸克甚至是顶夸克也会占有一席之地。只有那些质量小于所讨论能量尺度的夸克，才会作出重要的贡献。这就说明 Λ 的值依赖于有多少味的夸克参与该过程，这可以在强耦合参数和 Λ 值关系的因子 $33-2n_f$ 中体现出来。通常 n_f 是 4 或 5，相应的 Λ 值会有很大差别。

因此引用 Λ 值的时候会带有一个说明，表示它包括多少种夸克味道（是 4 个还是 5 个）。严格地讲，这是一个表示所用重正化

方式的标签。人们所知的 Λ 的值并不是很精确，约为 200 兆电子伏。对于 5 种类型的夸克并且使用理论物理学家们公认的习惯，QCD 理论中的“特征数”的表示方式为 Λ≈200 兆电子伏，其中“≈”代表“大致相等”。这看起来并不漂亮。

夸克变得更轻

一个夸克有多重？尽管这看起来是一个特别容易的问题，但实际上它的答案绝不是那么简单。夸克的质量和 QCD 的耦合“常数”一样狡猾，夸克的质量依赖于你怎样看待它，由于禁闭的存在，这也是一个棘手的问题。

由于夸克被限制在质子、π 子以及其他强子中，实验物理学家们无法奢望能够用测量电子的那种方法来测量自由夸克的质量。因为夸克永远不可能是自由的，夸克的质量只能通过它在某种相互作用中的效果体现出来。因此夸克的质量不是唯一的，而是依赖于它们被禁闭的方式。任何夸克质量值都应该注明脚标，用来表示得到该质量值的方法。

例如，其中一种确定夸克质量值的方法就是利用重子的夸克模型，它能够模拟重子质量产生的过程，通过将模型套用到强子质量的测量值上来确定夸克的质量值。因为每一个夸克都住在塞满相互作用着的夸克和胶子的强子“包”中，将它们粘在一起的是一种像粘糖浆一样的极强相互作用，它们的行为看起来就好像具有某种“等效质量”一样，这种等效质量并不一定需要和“真实”质量相等。这种等效质量被称作组分质量。粗略来讲，上夸克和下夸克的组分质量大约都是质子质量的 1/3，因此各自的质量大致都是 300 兆电子伏，奇异夸克的质量要大得多，其组分质量的值大约是 550 兆电子伏。

怎样让这些由夸克模型所得出的质量，同经常引用的其他一

些夸克质量值相一致呢？在第四章中，我们提到过上夸克和下夸克的质量分别为 4 兆电子伏和 7 兆电子伏，这与夸克的组分质量 300 兆电子伏左右大相径庭。这些较小的值是通过 π 介子衰变数据得到的，它们被称作通用夸克质量，是通过某种方法得出的。在这种方法中，很小但是有限的夸克质量会破坏一种对称性，即一套无质量 QCD 理论所具有的那种额外对称性。这一点还会在第九章进一步谈到。一个组分夸克实际上就是环绕着胶子云的通用夸克，正是这种胶子云携带了大部分质量，也正是它引起了这两套质量值的差别。

还不止这些。夸克的质量作为一个参数，在 QCD 理论的主方程中也极其重要，但这并不是理论预测到的。该质量参数并不是物理质量。两点间的格林函数，表示了夸克从一个地方平稳地移动到另一个地方的可能性，物理质量就是指在能给出这种可能性的完整表达中起重要作用的质量。这种质量还有一个名字——“极质量”（pole mass）。例如你能够测量电子的极质量，它就是在数据表中给出的电子质量的值，但夸克的极质量是无法测量的。最好的粒子物理学家可以在主方程中加入质量，在微扰论有效的高能区域对其重正化。这样夸克质量就会像依赖于能量的耦合参数一样，依赖于能量了。在这种方法中，夸克的质量并不固定，它变成了一种随着能量增加而减小的“跑动”能量。这种夸克质量的跑动像耦合常数的跑动一样，也是 QCD 理论的一部分。

从好多年前开始，理论物理学家们就已经理所当然地认为夸克质量是变化着的。但是这种夸克质量变化的第一次实验证据，是 1997 年在法国蒙彼利埃的一次 QCD 会议上，由利物浦大学的萨尔瓦多（Salvador Martii Garica）和他的同事公布的。他们利用在 LEP 的 DELPHI 上进行的 Z 粒子衰变为强子喷注的实验，搜集了两年多的数据。b 夸克所产生的喷注，包含了一些源自于最初电子-正电子相互作用点之外的粒子，这是 B 介子衰变相对

较慢的结果，见图 5.7。与之相反，你一般可以通过轻夸克产生的粒子喷注追踪到原始的相互作用点。因此在原始碰撞点之外起源的大量粒子，成为了 b 夸克的重要信号。这就使两种类型的喷注——b 夸克和轻夸克产生的喷注很容易区别。此外，b 夸克会产生更少的胶子，这是由于它们的质量较大。可以从产生 b 夸克喷注和轻夸克喷注的概率差别将其反映出来，因此这一差别也是测量 b 夸克质量的一种方式。萨尔瓦多和他的同事对比了两种喷注的概率，并且利用 QCD 理论，对电子-正电子碰撞过程中出现的三喷注事件进行了精确的计算。在那一年，他们发表了推导出来的 b 夸克质量值：2.67±0.50。这个数据是在 91 吉电子伏的能量下得到的，91 吉电子伏也是 Z 玻色子的质量。它只占从 ϒ

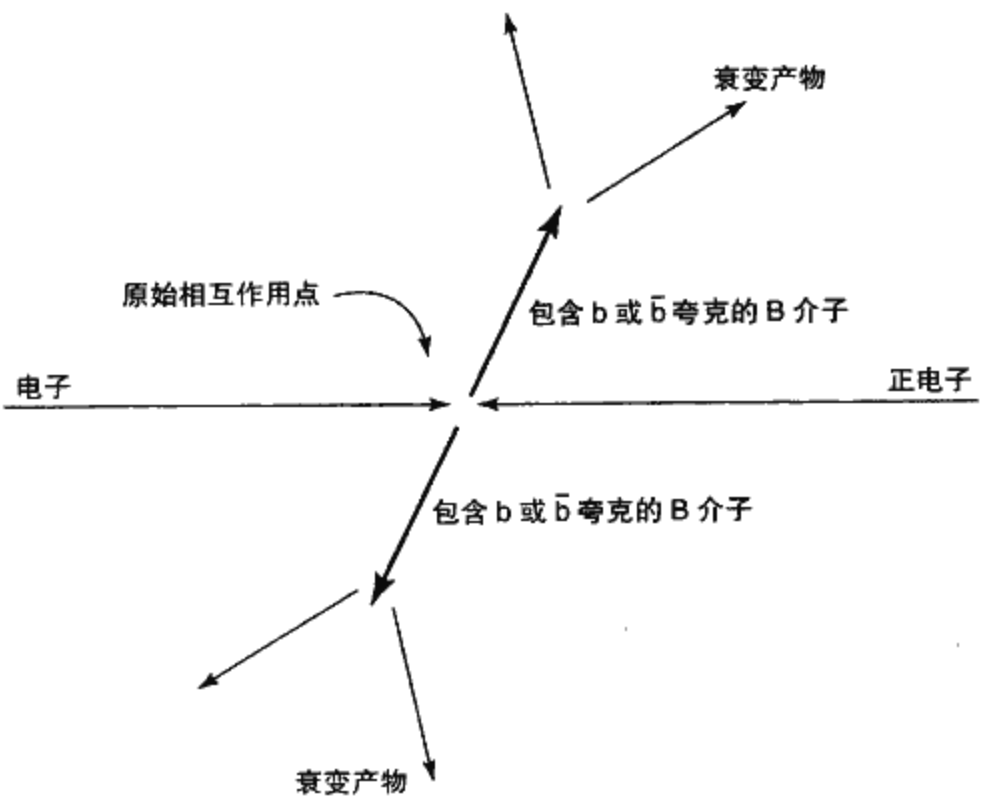


图 5.7 相对较慢的 B 介子衰变的粒子轨迹，这些轨迹的起始点离最初的反应点有一段距离

介子出发得到的 b 夸克质量值的 60%，而后者的能量是前者的 1/20。b 夸克的质量随着能量增加而减小，这符合 QCD 理论关于跑动质量的预言。

第六章 群居的胶子

物理学可以是一个又脏又累的行当。一方面是优雅的数学理论、实验室中闪亮的真空室、闪闪发光的电子以及精确的计算；另一方面却是真实、神秘而又凌乱的世界。许多实际的物理系统，仍然无法得到完满的解释，从错综复杂的恒星运行到从桶里倒出的沙流，这种系统的数量不计其数。但有时，实验物理学家们也能够发现一些明显的证据，从而揭示出大自然的简洁本质。在粒子物理学中，见证这种简洁性的一个好地方，就是粒子喷注现象。图 6.1 中显示了一次粒子喷注事件。

喷注和瑞士钟表

一个在 Tevatron（费米实验室正负质子对撞机）环中飞速运动的质子，迎面撞上了一个反质子。从理论上讲，以这种方式产生的粒子会朝着各个方向随机运动，但实际不是这样。除了质子和反质子本身的残骸之外，粒子还会以喷射出两个、三个或四个圆锥形的粒子束的形式出现。在 LEP（大型正负电子对撞机）或者其他电子-正电子对撞机中，也会产生相似的现象。与质子和反质子等复合粒子对撞过程相比，该过程更为引人注目，因为参

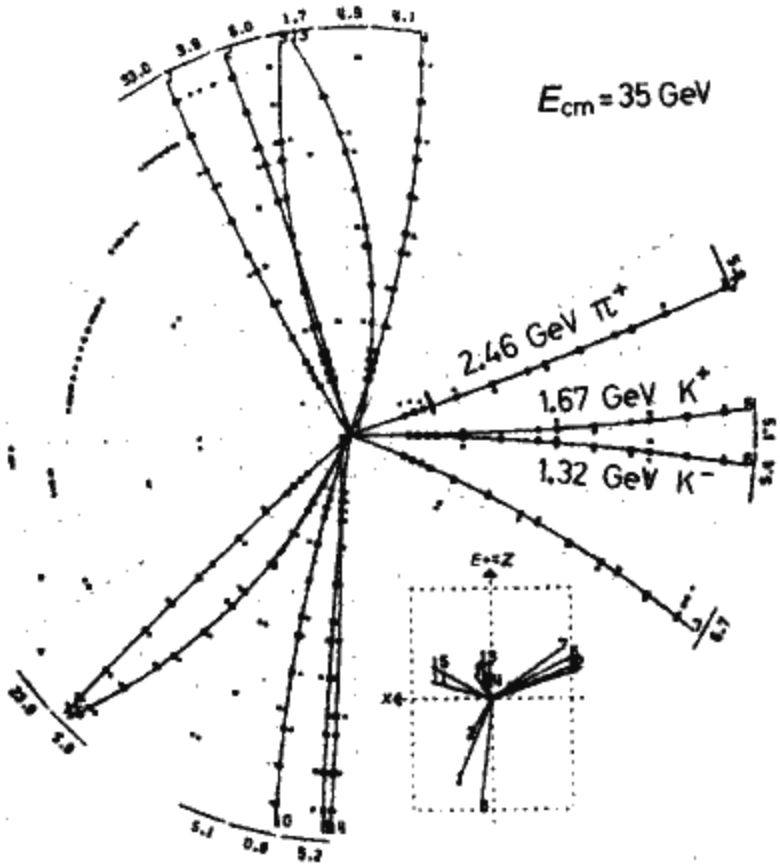


图 6.1　1979 年在正负电子环状队列加速器（PETRA）上发生的一次三喷注事件，它为胶子的存在提供了证据。图片来源：德国汉堡，德国电子同步加速器（DESY）

与该过程的粒子都是基本粒子。这些产生的粒子流，就是此前已经在第四章和第五章中遇到过的“喷注”。

最原始的喷注事件，是电子-正电子对湮灭，生成两束强相互作用粒子喷注的过程。在该过程中首先会产生夸克-反夸克对，随后每一个粒子相应地又会产生一个强子喷注，见图 6.2。这就是喷注最吸引人的性质之一，人们立即将能解释喷注形成的简单图像看做相关费曼图的核心内容。喷注本身不仅是实验者们在探测器中观测夸克和胶子轨迹的最直接方法，也可以用来研究基本的夸克和胶子过程。在电子-正电子碰撞中出现的多喷注事件，

是现今粒子物理学图像的精髓。

当然，也存在一些问题。那些最初产生的裸夸克和裸胶子是怎样自己穿上衣服，从而变成了实验者们所观测到的粒子的？我们至少需要一个从根本上解释这些过程的可靠模型，否则我们就无法将计算结果同实验者实际测量到的粒子数值作比较。

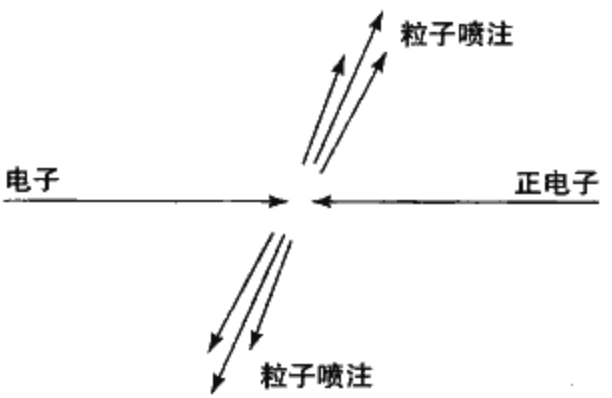


图 6.2 典型的喷注事件：一个电子和一个正电子发生湮灭，产生两个粒子喷注

即使原始的夸克或反夸克想要独立出来，它们也不能通过直接的方式来完成。这是因为自由夸克和反夸克是不允许存在的。禁闭规则要求，夸克和反夸克都应该束缚在复合粒子当中。那些从碰撞中产生、试图向四面八方飞出去的夸克和反夸克也是这样，禁闭规则会坚持它的要求。由于夸克和反夸克对有分离的趋势，从能量角度考虑，最可能的情况是每一个粒子都能够从真空中再揪出一个粒子，与之形成新的夸克-反夸克对。那些从碰撞点飞散开的夸克就是这样，它会同从周围真空中产生的夸克-反夸克对中的反夸克结合，产生一个介子，见图 6.3。这样就会留下一个新产生的夸克，它反过来又会激发另一个夸克-反夸克对，然后再同其中的反夸克结合形成介子。结果会导致一长串介子链的产生，每一个介子的能量都要比前一个小，因为能量在产生夸克-反夸克对的过程中被吸收了。当最后一个夸克没有足够的能

量从真空中激发另一个夸克-反夸克对时，这条链就中止了；在这一点上，该夸克会同反夸克介子生成链中的一个同样“虚弱”的反夸克结合。这使色荷得到平衡，最终产物不会带有剩余的净色荷。夸克和胶子变为强子的过程被称作强子化（hadronization），或者称作裂化（fragmentation）。在真实散射截面计算中，这是尚未被发现的最后一环。

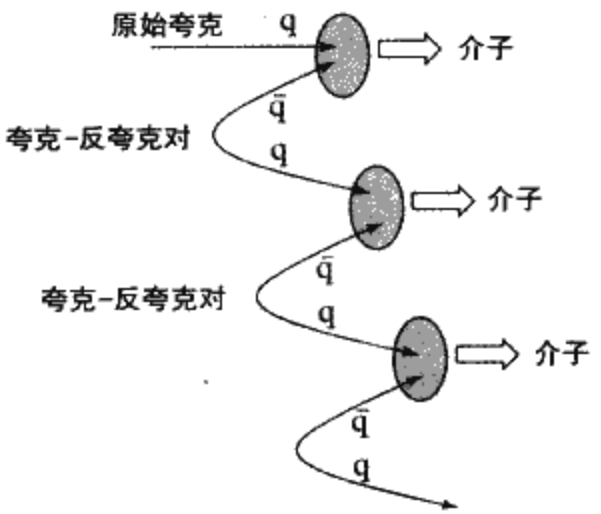


图 6.3　夸克碎片。一个原始夸克从真空中激发出一个夸克-反夸克对，原始夸克同反夸克结合为可观测到的介子。新产生的夸克再次重复该过程

在裂化期间，一个夸克会将它的动量分配给新的介子和子夸克。这可以用一种叫做裂化函数的概率分布来表示，裂化函数可以给出父夸克传递给介子的那部分动量。裂化函数是一种相反的结构函数：结构函数描述了在核子中找到具有某种动量的夸克和胶子的可能性；而裂化函数表示了给定父夸克或胶子，能够产生具有特定能量的介子的可能性。同结构函数一样，想要通过基本原理计算裂化函数，物理学家还需要知道更多低能状态下，非微扰 QCD 问题的处理办法。

裂化同样也吸引了大批学者投入到解读这种现象的事业中

来。正如上面所提到的，他们通过加入重子产生以及胶子分裂为夸克-反夸克对的过程，很快扩展了简单但存在固有缺陷的“独立喷注模型”。人们利用改进后的模型，可以理解质子-反质子碰撞的数据。另一个被广泛应用的模型是集团裂化模型（cluster fragmentation model），该模型是指辐射胶子会生成夸克-反夸克对，它们会和其他夸克和反夸克（不必来自同一个父胶子）形成不带净色荷的粒子集团。接着，这些由夸克、反夸克以及胶子组成的色单态粒子集团会变成可观测的粒子喷注，主要是 π 介子和非常少的奇异介子。第三种著名的模型是在夸克和反夸克飞散的时候，把它们之间的色相互作用看做是一个狭长的管道，或者称之为“弦”。在这种“弦模型”中，弦被撕裂成许多碎块，变成了粒子。这有点儿像一个带南北极的条形磁铁，把磁铁砸碎后就会变成两块条形磁铁，每一块都有各自的南极和北极。胶子会在弦上产生结，表现为最终探测到的粒子角分布。这种方法可以非常满意地解释电子-正电子碰撞三喷注事件的粒子角分布。但强子化理论仍然是一团乱麻——一方面，粒子物理学家们不得不依靠强子化模型；另一方面，这些模型在理解粒子喷注的物理过程中总会有局限。

按照麻省理工学院的罗伯特·贾菲（Robert Jaffe）的解释，强子化就是让质子和反质子发生碰撞，然后研究碰撞碎片，该过程就像把两部瑞士钟表互相撞碎，最后却只能找到更多的钟表，而不是齿轮、发条以及其他零件。同样，粒子物理学家们也没有发现夸克和胶子，看到的都是完整的复合粒子——也就是更多的瑞士钟表。

喷注的世界

歌曲《西区故事》（*West Side Story*）中唱道：“如果你是一

束喷注，那么你将会永远都是喷注。”① 问题在于，看到确实是喷注的喷注并不总是那么容易的事情。我们不去管它，先数一下喷注的实际数量。例如，某些碰撞所产生的粒子分布可能是两束大的喷注，也可能是一束大喷注和两束小喷注，这些都是极为普通的情况。但这种差别很重要，因为人们可以根据喷注数量提供的线索，直接窥探相互作用的核心过程。

结果造成出现一大堆听起来很奇怪的物理量，研究者们经过多年时间提出了这些物理量，试图将“喷注现象”（jetiness）量化，将各个喷注区分开。例如将每一个观测到的粒子的动量结合称为球度（sphericity），如果所有粒子的动量都沿着同一个轴方向排列，球度的值就为0，也就是说，这是一次完全的二喷注事件；如果产生的粒子在各个方向上都是平均的，那么球度的值就是1。球度最初是由布鲁斯基和比约肯提出来的，早在1975年，在SLAC的SPEAR电子-正电子对撞实验中，MARK I的实验者们就用这个参数来约束第一束夸克喷注。与球度相反的是“绕冲度”（thrust），它的值不是根据垂直于选定轴的粒子动量得出，而是平行于该轴：一个背靠背的二喷注事件，其绕冲度值为1，球对称事件的绕冲度值为0。球度和绕冲度都是用来判断一个事件在整体上是否是一次喷注事件的工具，研究者们还提出了其他方法，来区分一束大喷注和两束小喷注。一种识别喷注的方法就是看从碰撞点射出的能量，在某个圆锥角内的分布是否大于某个特定值，在实验中可以用探测器中的量热计来测量该能量。另一种方式是发明一种与能量和飞行方向相关的量，它可以适用于每一对粒子。具有最小值的粒子对被看做是粒子集团，然后重复该过程，将这个粒子集团当做单独的粒子来处理。最终，经过足够

① 译者注：从文字上，原文可以那么理解，但是原文中jet是一个帮会名，因此这句话中jet本来是jet帮会成员的意思。

多次的重复计算，所有粒子都分配给了粒子集团，这些粒子集团最终形成了喷注。

电子-正电子碰撞是一个观测喷注的好地方。一个电子和一个正电子湮灭会产生一个虚光子（或者在足够高的能量下产生一个 Z 粒子），它能够衰变为夸克-反夸克对。这些夸克-反夸克对又能物质化，成为两束背对背的强子喷注，它们的方向能够反映出原始夸克-反夸克对的方向，实验物理学家们能够实际观测到这些强子喷注。并且，随着碰撞能量的提高，实验者们观测到的喷注也变得越来越像喷注了；换句话说，也就是每束粒子相对于总粒子束方向的相对分布减小，因此每束喷注会更加明显。这在人们的预料之中。

严格校准的喷注，允许实验者们测量每束喷注相对于碰撞粒子束的偏转角。如果喷注是由夸克和反夸克形成的，因为它们都像 μ 子和反 μ 子一样具有 1/2 自旋，其角度分布会同通过电子-正电子湮灭产生 μ 子和反 μ 子的过程一样，依赖于自旋。SLAC 在 1975 年得到的结果证实了这一点，从而为存在自旋为 1/2 的夸克提供了有力的证据。

在足够的能量下，下一层复杂性会显现出来，产生的夸克或者反夸克会发射胶子。结果会产生三喷注事件，CERN 的理论物理学家约翰·埃利斯（John Ellis）、玛丽·杰拉德（Mary Caillard）和格雷厄姆·罗斯（Graham Ross）在 1976 年认识到了这一点。这种三叉形事件于 1979 年在 DESY 的 PETRA 存储环上第一次被观测到（图 6.1）。这是第一次观测到有关胶子的证据，它标志着实验上发现了胶子，尽管随后 PETRA 的四个小组在谁首先发现了该结果的问题上很不体面地争论不休。

与二喷注实验相比，计算三喷注事件发生的概率，要涉及额外出现的强耦合参数，这同新出现的夸克-胶子-夸克结点吻合，见图 6.4。我们先把复杂的因素抛在一边，二喷注事件的散射截

面同三喷注事件的散射界面的比值会给出强耦合参数的值，它同理论的不确定性无关，这是 SLC 和 LEP 的实验者用来测量强相互作用强度的一种方法。三喷注实验的数据还显示出胶子的自旋值为 1，预言胶子自旋为 1 的理论令人信服地同实验完全一致，而胶子自旋为 0 的想法很明确地被剔除掉了。

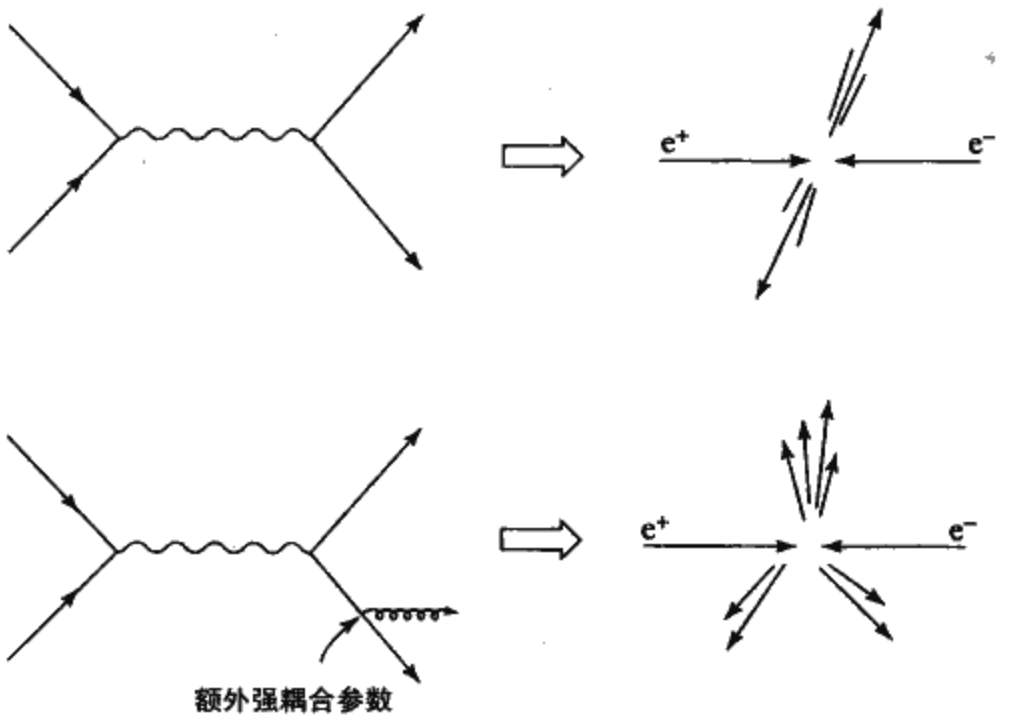


图 6.4　二喷注事件和三喷注事件

随着三喷注甚至是多喷注事件的发生，我们有可能将夸克和反夸克引起的喷注和那些起源于胶子的喷注区别开。这也同理论预计的情况相符，实验者们发现胶子喷注包含更多的粒子，并且比夸克喷注更粗，这是由于胶子带有更大的有效色荷，因而能产生更多的胶子辐射。人们还发现胶子喷注的总电荷为 0，而夸克喷注的电荷值却不是 0。

区分夸克喷注和胶子喷注的另一个线索是：在三喷注事件中，夸克喷注和反夸克喷注形成的夹角方向中存在的粒子数目，

比这两束喷注同胶子喷注形成的夹角方向中的粒子数目要少，见图 6.5 中的图（a）。这种现象是由一种能引起分裂的干涉效应造成的，这种干涉效应即所谓的色相干性，能够引起“软”胶子（即低能胶子）从夸克或者反夸克中发射出来。色相干性会使夸克中发出的胶子同夸克运动方向的夹角越来越小，见图 6.5 中的图（b），因此能给出一个角度序列，这同 QDE 理论中发射光子的情况相对应。

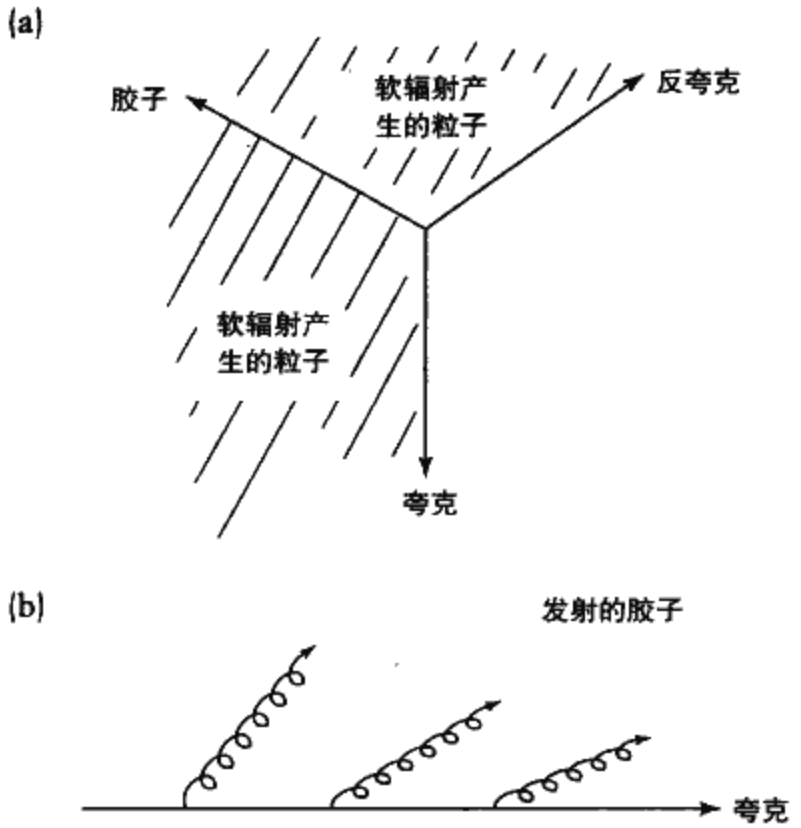


图 6.5 色荷相干性：沿着粒子束轴看过去，电子-正电子三喷注事件有软辐射粒子区，这是由色荷相干性决定的［图（a）］——产生角度序列的干涉效应［图（b）］

如果碰撞的能量更高，就能够完成一项检验 QCD 理论的重要实验，KEK 的 TRISTAN 首先达到了这个要求，之后 CERN

的 LEP 和 SLAC 的 SLC 也分别达到了这一能量。整个理论结构最关键的部分，就是它的非阿贝尔性质，这会导致胶子和胶子的相互作用；与之不同的是，在阿贝尔性的 QED 理论中，光子和其他光子之间不会存在相互作用。换句话说，QCD 理论预言存在着三胶子顶点。要证明这一点，就要测量四喷注事件，然后就能够在那些“更加平凡的”原始背景中数出包含三胶子顶点事件的个数，见图 6.6。1989 年，在 KEK 的 AMY 合作小组第一次发现了包含胶子自耦合的四喷注事件。到 20 世纪 90 年代初，LEP 已经能够大量生产三胶子顶点事件，其产量足以用来进行细致的研究，它们同那些有关色荷群的直接证据都已经在第四章中讨论过了。发现三胶子顶点的秘密在于对喷注间夹角的仔细研究，因为能够产生另外两个胶子的那个胶子具有与众不同的自旋组分，这在最终出现喷注的角分布中可以表现出来。

电子-正电子并没有完全垄断喷注市场。根据 QCD 理论，一

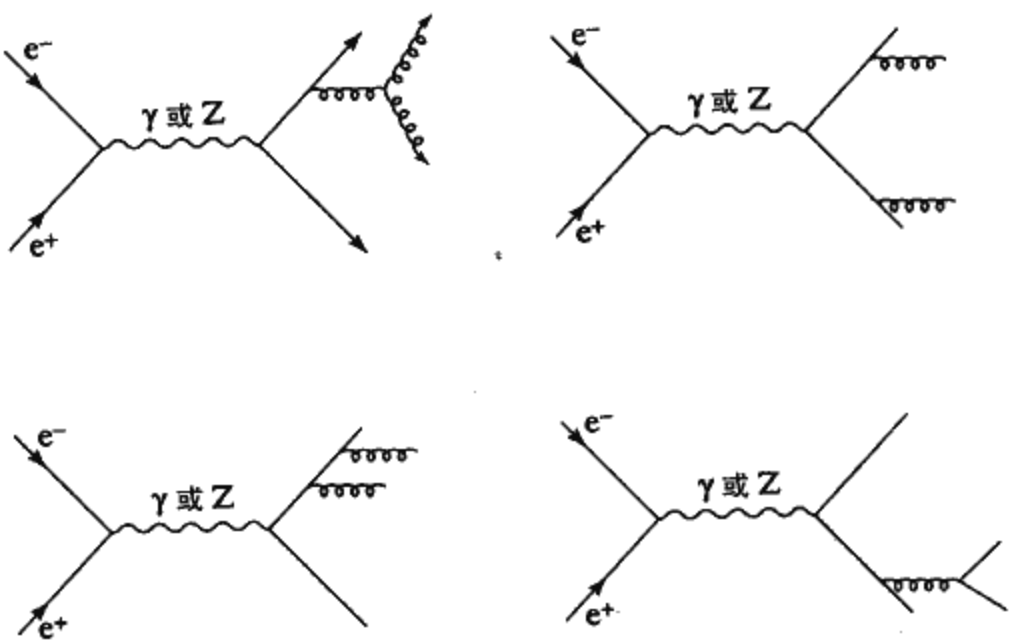
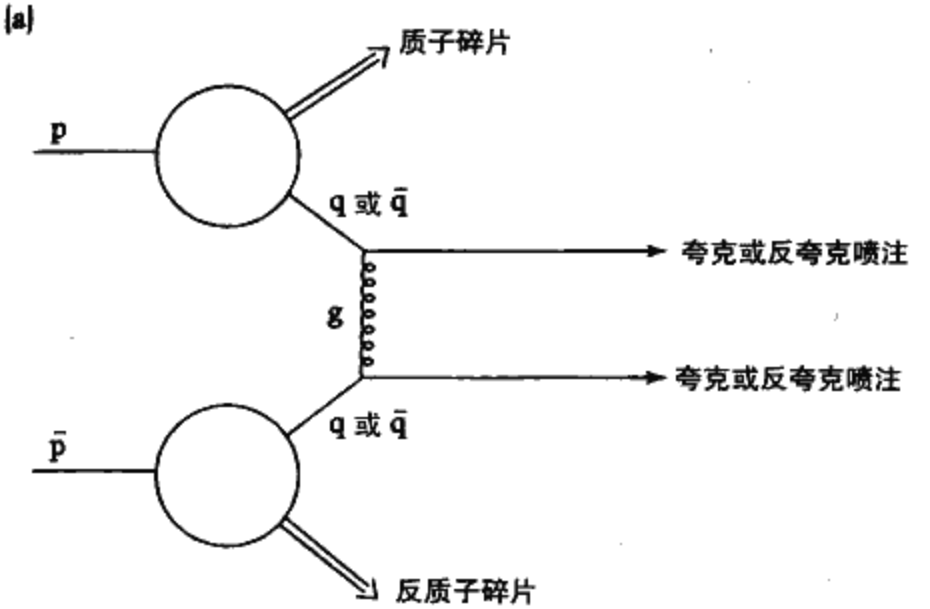


图 6.6　在电子-正电子湮灭中的四喷注事件

个质子和一个反质子的碰撞过程，实际上是通过从质子中的夸克或反夸克与反质子中的夸克或反夸克交换胶子而实现的，这种过程最终也会导致喷注的产生，见图 6.7 中的图（a）。另一种产生



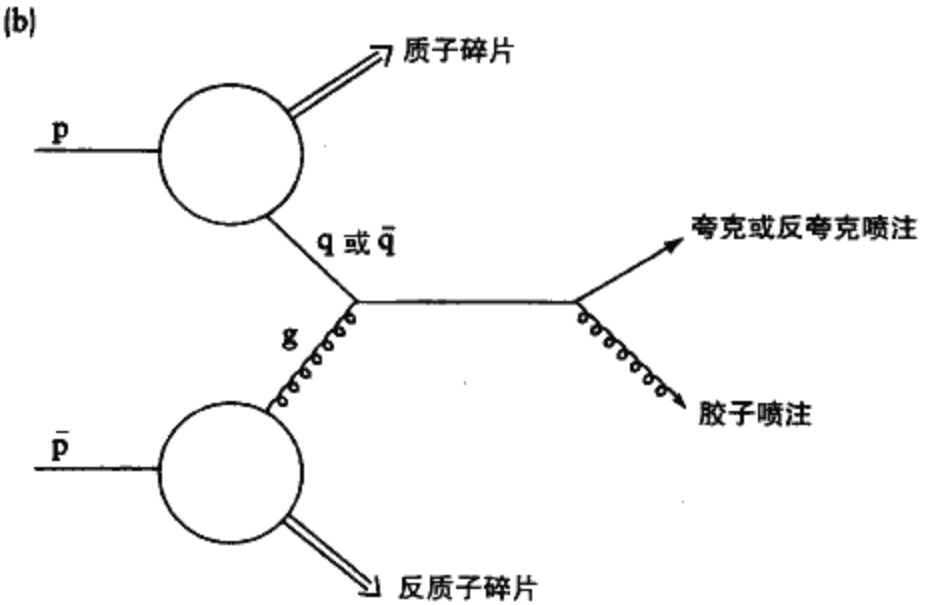


图 6.7　质子-反质子碰撞产生的二喷注事件，图中展示了胶子之间的散射［图（a）］和夸克（或反夸克）同胶子的散射过程［图（b）］

喷注的方式，是一个粒子产生的夸克和另一个粒子产生的胶子散射，见图 6.7 中的图（b）；这种过程存在许多可能的模式，其中一些过程还包括三胶子或四胶子顶点。但是只有在喷注具有很高的能量以及喷注相对于入射粒子束的夹角很大的情况下，才能够很清楚地看到这种喷注事件。直到 1982 年，人们才发现这种喷注存在的明确证据。这是由 UA2 合作小组发现的，该小组当时利用新建成的 CERN 的 $Sp\bar{p}S$ 机器展开研究。在 Tevatron 更高的能量下，喷注可以被显示得更清晰，许多事件都具有两个、三个、四个甚至更多的喷注。在 Tevatron 上观测到的二喷注事件的数量以及它们的能量和角度分布，还有大量其他喷注散射截面，都支持了 QCD 理论的预言。

但是，强子-强子喷注确实带来了特殊的挑战，因为这些参与碰撞的粒子在效果上相当于由夸克、反夸克以及胶子所组成的包，而不是简单的基本粒子。总的碰撞概率主要是由相对较弱的碰撞构成，只有那些最猛烈的碰撞才能给出研究者们所谓的“硬散射”现象。首先，它能够产生相对于入射粒子束的大角度粒子喷注；其次，它与基本的夸克、胶子散射过程有简单的关系。粒子束喷注包含质子同反质子碰撞碎裂后的复杂碎片，它们使理解碰撞碎片变成夸克和胶子喷注的过程更加混乱。还存在另一个问题，计算碰撞概率，就意味着要考虑质子和反质子中各种夸克、反夸克和胶子贡献的可能性；换句话说，就是要利用第四章中已经引入的结构函数。这是一个全新的潘多拉盒子，我们将在第八章中把它打开。

胶子的玩笑

任何喷注都会包含胶子辐射。考虑所有的辐射胶子，我们就可以看出 QCD 理论的精妙之处和瑕疵。如果从夸克中辐射出的

胶子同夸克运动的方向相同，也就是两者共线，就会出现一些问题。另一种与之相关的情况是辐射出的胶子很弱，也就是很“软”，它的能量可以等效为零。这两种表面上平淡无奇的情况，都会导致微扰场论中最令人讨厌的结果——无穷大。因为这两种无穷大都是在处理低能问题时产生的，因此它们都是被理论物理学家称为“红外发散”的例子。共线的胶子和软胶子都预示计算结果会无穷大。在这两种情况下，问题的根源都在于胶子没有质量。

令人难以置信的是，这些无穷大会和另一种隐藏在计算过程中的无穷大抵消。电子-正电子散射生成夸克和反夸克的过程可表示为 $e^+ + e^- \to q + \bar{q}$，图 6.8 中的图（a）是描述该过程最简单的费曼图。要进行更精确的计算，就要用到更复杂的费曼图，图中会含有一些包含胶子的封闭圈，见图 6.8 中的图（b）。这些封闭圈曾经以紫外发散的形式给我们制造过麻烦，但可以通过重正化方法加以解决。但它们也在能量谱的另一端遗留下了尚未解决的问题：在圈中的动量降到零时，同样是这些圈也会造成无穷大，这个问题也起源于胶子没有质量。这些与软胶子和共线胶子相关的无穷大都具有共同的原因，这就为抵消这两种灾难铺平了道路：把胶子圈图［图 6.8 中的图（b）］和真实的胶子辐射图［图 6.8 中的图（c）］的贡献加到一起，红外发散就能够被消除。

这是天方夜谭吗？圈图适用于过程 $e^+ + e^- \to q + \bar{q}$，而胶子发射图明显适用于不同的过程 $e^+ + e^- \to q + \bar{q} + g$。因为在真实实验中，软胶子和共线胶子变为粒子的能力都很弱，解决上述矛盾的一种方案，就是认为过程 $e^+ + e^- \to q + \bar{q} + g$ 和过程 $e^+ + e^- \to q + \bar{q}$ 完全无法区分。消除红外发散的关键是要将所有不可区分的末态加起来，这种方法也可以用来处理更加复杂的费曼图结合问题。利用它可以相应地剔除微扰论中人为引入的部分。在 QED 理论中就已经存在的这个问题，确实可以解决——实际上，QCD

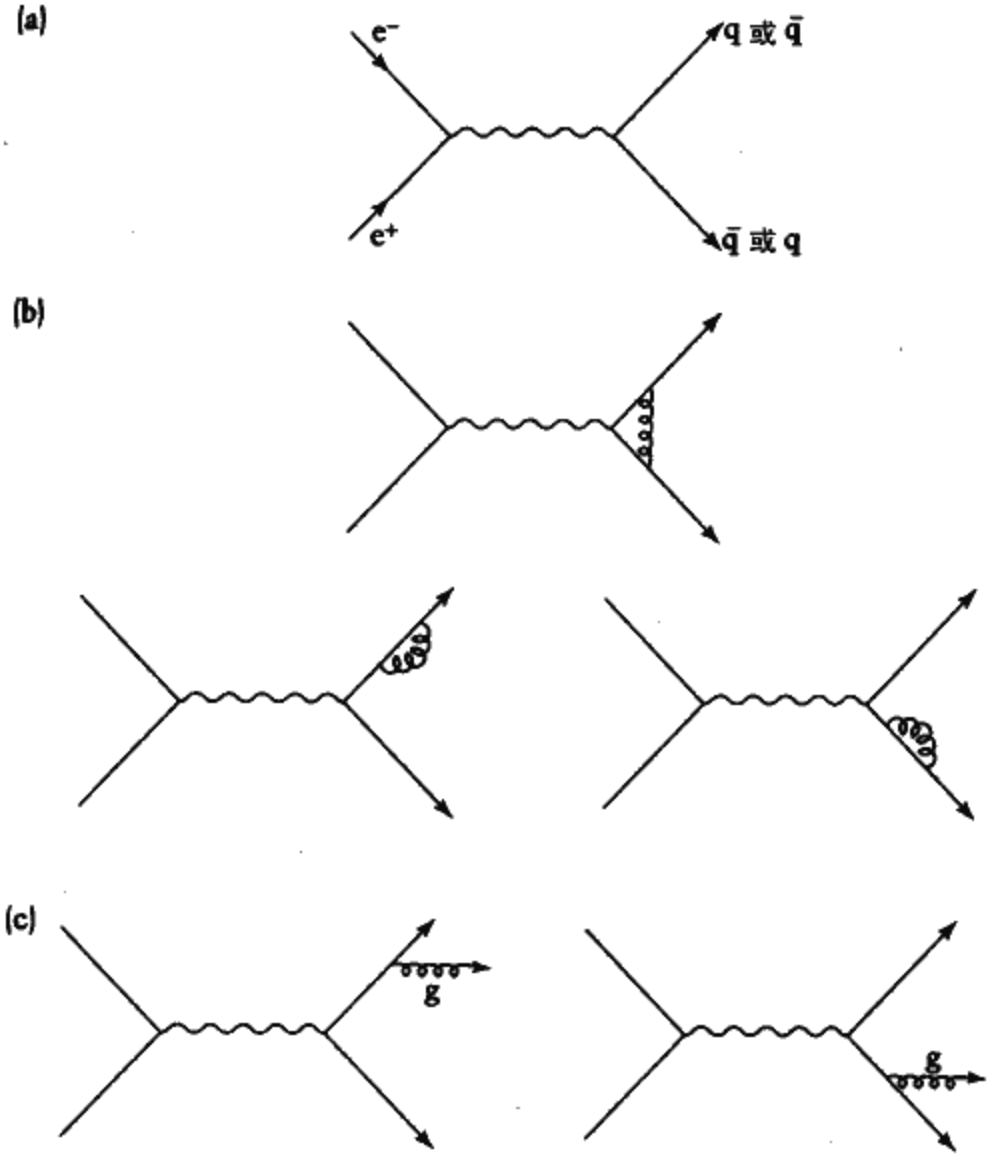


图 6.8　在电子-正电子湮灭过程中发射胶子的简单过程

理论中的所有计算都受到红外发散的潜在威胁。因此，如果不将这一部分抵消，理论物理学家们的计算就没有任何意义。

在 $e^+ + e^- \to q + \bar{q}$ 的计算过程中，还有一个巧妙之处，这在 QCD 理论中也是很罕见的一次“免费午餐”。要完整地计算该过程的散射截面，就意味着要计算无穷多个费曼图，最开始的几个就是图 6.8 所示的那些费曼图，接下来的费曼图复杂程度会越来

越高。但是最简单的圈图和它们很相似，在 QED 理论中讨论重正化时也出现了相应的费曼图。实际上，QCD 圈图是一整套费曼图的一部分，这套费曼图是理解能量增加与 QCD 耦合强度降低关系的关键。因此，如果在一次简单的散射截面计算中，理论物理学家们使用跑动耦合参数来代替固定耦合参数，他们马上就能够将一整串相关的费曼图合并起来。实际上，他们对所有可能的费曼图的一个子集进行了求和，这个子集本身也具有无穷多的项。这就是重正化的效果：仅仅对自由费曼图求和，就正好可以消除散射截面计算对重正化尺度的依赖。

但是，散射截面或者人们所研究的其他任何过程，都没有计算和加进所有的费曼图。为了进一步研究，人们必须对有更多圈的费曼图进行直接研究，以得到跑动耦合常数的更精确表达式。在复杂程度更高的散射截面计算中，这种更精确的描述能够考虑更多的无穷项。计算的圈图越多，就会产生越精确的耦合常数表达式，相应地就能计算更多的项。因此，这虽然是免费的午餐，但分量并不是很足。这更像是我们对已知方法的更有效应用。

胶球和其他新奇的事物

胶子之间存在相互作用。但是有什么东西使它们无法聚集到一起，结合成一个只包含胶子的粒子团呢？根据理论物理学家们的解释，答案并不存在——胶子能够结合到一起，形成一种新的中性介子，称作胶球（glueball）。在高能物理中，胶球与其他粒子相比，可能已经被“发现”好多次了。

一只不在纽约的胶球

在位于纽约州伊萨卡市的 CESR 电子-正电子对撞机上，实验物理学家们研究了一种重介子的性质，这种介子有一个难听的

名字，叫做 $f_J(2200)$。这是一种存在问题的介子。第一次发现该粒子的时间可以追溯到 1986 年，在 J/ψ 衰变中出现了一种寿命相对较长的态，它的真实性从一开始就受到了质疑，但它的地位最终逐渐巩固起来，变得“确定无疑”。它的自旋很可能是 2，但也可能是 4，没有人敢确定；它的一些衰变路径只是根据实验列了出来，而不是测量到的。物理学家们感兴趣的是这种假想粒子诞生和消亡方式之间的关系。

在 CESR，CLEO 的实验工作者们仔细检查了碰撞产物，试图寻找 π 介子对，每一个 π 介子对都包含一个正 π 介子和一个负 π 介子。$f_J(2200)$ 介子就能够衰变出一组这样的 π 介子对，π 介子对的总能量与 $f_J(2200)$ 介子的质量相符，π 介子对数量出现峰值就表示在电子-正电子湮灭过程中产生了 $f_J(2200)$ 介子。尽管实验物理学家在实验中发现了 π 介子，但还是没有发现 $f_J(2200)$ 介子存在的任何迹象。

事实证明，这些结果对于那些寻找胶球的人来说是一个好消息。因为一个由带电夸克和反夸克所组成的常规介子，不仅能够通过在带电荷的组分之间交换电子产生电磁相互作用，还能够通过在带色荷的夸克和反夸克之间交换胶子产生强相互作用。与之相反的是，胶球是由胶子组成的集团，不包含任何带电荷的组分，因此对光子不敏感，但它能够感受到所有强相互作用的影响。这种强相互作用和电磁相互作用之间的失衡是用“黏度”(stickiness) 来衡量的，它表示胶子相互作用强度与光子相互作用强度之间的比率：黏度值大，就说明对电磁相互作用反应弱，从而说明存在胶球。

CLEO 的结果给出了 $f_J(2200)$ 发生电磁相互作用的限制。为了得到黏度的数值，CLEO 的实验物理学家们将他们的结果同北京高能物理所（the Institute of High Energy Physics，简称 BHEP）那台 4.4 吉电子伏的北京电子-正电子对撞机（Beijing

Electron-Positron Collider，简称 BEPC）所得到的结果综合到了一起。北京谱仪组在 BEPC 上观测到了 $f_J(2200)$ 介子，但是是通过 J/ψ 粒子衰变为一个光子和一个 $f_J(2200)$ 介子的过程得到的，$f_J(2200)$ 介子最终又衰变成一个 π 介子对。通过衰变链的最后一环可以测量出 $f_J(2200)$ 介子和胶子之间相互作用的强度。把强度值带入到黏度的计算中，复合结果给出的黏度值很高，这同人们预计胶球的性质相符。这些复合结果还显示，$f_J(2200)$ 介子在电子-正电子碰撞中通过 J/ψ 粒子产生，而不是直接产生，这说明它同普通夸克-反夸克对组成的物体属于不同的类型。其他的一些性质也说明这种粒子的与众不同：人们从没有在质子-反质子对碰撞过程中观测到它，但 BES 合作小组却观测到了 $f_J(2200)$ 介子衰变为质子-反质子对的过程。在 LEP 观测到的许多依赖于胶子喷注的三喷注事件中出现了一种现象，它很可能是 $f_J(2200)$ 介子出现的证据，$f_J(2200)$ 介子在二喷注事件中不会出现。

在计算机的支持下，QCD 理论很快指出，模拟自旋为 2 的胶球所得到的质量十分接近 $f_J(2200)$ 介子的质量。将这个结果和实验得到的证据联系起来，你就会发现 $f_J(2200)$ 介子越来越有可能是一个胶球。但到目前为止，这一切都是不确定的，$f_J(2200)$ 介子只能被当做是胶球的一个候选者。它是否真的是胶球仍然是一个需要澄清的问题；这也说明了在科学中，并不是所有问题都像有时候那样黑白分明。经过了多年的搜寻，虽然人们已经多次公布发现了胶球，但粒子物理学家们还是不能将任何胶球的候选者确认为真正的胶球。例如，存在一种自旋值为 0 的 $f_0(1500)$ 介子，它的胶球地位比 $f_J(2200)$ 介子还要稳固，在 1996 年它成为了胶球的第一候选者。那时，物理学家们认识到，过去的数据显示 $f_0(1500)$ 介子同样是 J/ψ 粒子衰变的产物，它的黏度值更符合胶球的要求：一条相关报道的标题写道“过去被冷落的粒子符合胶球的特征”。

但是还存在着难题——到底什么才是“胶球的特征”？换句话说，当一个胶球经过的时候，实验物理学家们期望在探测器中观测到什么？它不带电荷，这使得我们很难看到它。一些在许多粒子反应中起指导作用的守恒定律对它都不起作用，估计它的质量要依赖于计算机模拟。如果你觉得这些性质还不算糟糕透顶的话，尚未理清和确定的不同介子态也复杂得难以置信，尤其是胶球情形，介子态能够和具有同样量子数的其他介子态混杂到一起。搜寻胶子的人，至多只能期望能观测到一种由胶球和至少另外一种粒子产生的复合信号。

奇异的发现

在纽约州另一面的长岛，布鲁克海文国家实验室 AGS 质子加速器的实验者们，着手研究了胶球和“常规”介子之间的模糊区域。他们除了发现一个“候选者”之外，还发现了别的东西。

在编号为 E852 的实验中，高速质子轰击靶，通过强相互作用反应生成 π 介子。用一套磁铁系统将这些 π 介子引出，让它们打到液氢上。π 介子同液氢中的质子碰撞，会发生许多反应。实验 E852 中最吸引人的是反应 $\pi^- + p \rightarrow \eta + \pi^- + p$，实验小组可以通过 η 介子衰变为光子对来确认该反应。综合 η 介子的质量和能量以及最终得到的 π 介子，在观察到的事件中会出现一个峰值。这个共振峰标志着存在一种质量大约为 1300 吉电子伏的中间粒子，该粒子随后衰变为一个 η 粒子和一个 π 介子。这是一种叫做 $a_2(1320)$ 的介子，是 π 介子的一个亲戚，在粒子计数率的峰值中贡献最大。

但这并不是故事的结局，因为 η 介子的角分布是不均匀的，如果有另一个具有不同自旋和角动量的粒子伴随 $a_2(1320)$ 一同产生的话，这种现象就能够得到很好的解释。能发现 $\pi_1(1400)$ 这种新粒子的确十分令人激动，因为在已有可靠的夸克模型中，

它是没有意义的。它是另外一种物质，是一种奇异物质。E852小组在1997年发表的论文中写道，“我们得到了这样一种结论，存在可靠的证据表明能够产生一种……奇异介子”。此后不到1年，CERN的晶桶合作组（Crystal Barrel Collaboration）就将上文中的“可靠”两字换成了“直接”。他们的探测器是一个桶形的阵列，它具有1380个光敏碘化铯晶体。他们利用CERN的LEAR反质子加速器让质子同氘靶相撞，主要研究反应 $\bar{p}+n\rightarrow\eta+\pi^{-}+\pi^{0}$，从而找到了通过 $\eta-\pi$ 共振子产生 $\pi_1(1400)$ 粒子的明确证据。第一种奇异介子①被证实了。

很明显，夸克模型允许存在夸克-反夸克所组成的介子，但这种介子的组分却并非如此。夸克模型包含了夸克自旋以及夸克与反夸克之间的相对运动，它预言存在一系列具有特定角动量和对称性的介子。但在引入了胶子之后，就允许存在一种“混血”介子。比如 $q\bar{q}g$——这是一种新的夸克-反夸克-胶子态，这些态改变了以往的规则，并且具有“奇异”性质，这些内容都是常规夸克模型所禁止的，但是它们却能帮助实验物理学家们搜索这些粒子。这种所谓的口袋模型在第七章还会再次进行讨论，它预言会存在一种质量大约为1400兆电子伏的混合粒子，这非常接近于 $\pi_1(1400)$ 粒子的质量——计算机模拟得到的质量要更大一些。还存在其他的奇异可能性，例如像 $qq\bar{q}\bar{q}$ 这样的四夸克态，或者看做是四夸克组成的介子-介子结合体 $q\bar{q}-q\bar{q}$。$\pi_1(1400)$ 粒子一开始被叫做 $\tilde{\rho}$（1405），人们假设它是一种 $q\bar{q}g$ 混杂，这是拥有正确性质的最简单情况。不久，另一种相关介子 $\pi_1(1600)$ 也出现了。

不久前，人们在2003年发现了一种由5个夸克组成的奇异粒子。这种东西被称作 Θ^{+}，它的质量是1540兆电子伏，是由2个

① 译者注：又叫混杂介子。

上夸克、2 个下夸克和 1 个反奇异夸克组成的，表示为 $uudd\bar{s}$。早在 1997 年，俄罗斯理论物理学家德米特里·迪亚科诺夫（Dmitri Diakonov）和他的同事就预言了这种粒子。人们最终发现，它存在的证据深藏在几个实验小组的有效数据当中。如果有人告诉他们在哪里寻找这些东西，这项耗时超过 35 年搜寻“五夸克”的工作会完成得更早一些。

奇特介子和多夸克态的发现以及许多研究者一致认为存在的胶球都表示，存在着全新的物质形式。这给理论物理学家们提出了新的挑战，因为这些物质都已经超越了“普通”夸克模型的界限。但即使抛开奇异物质，仅仅在 QCD 理论的框架内理解最简单的粒子，也是一项艰巨的任务。

第七章 夸克和强子

粗略地讲，原子是由一个带正电荷的核心，和环绕在核心周围的一团带负电荷的电子组成。原子核所带的电荷来源于质子，质子的数量等于轨道上的电子数目，这就得到一个很明显但十分重要的结论：原子总体上是呈电中性的。如果去掉一个电子，就剩下了一个带正电的离子。从原子中剥夺一个电子是很容易的，用热、光或者电离等方法都能够实现；离子是很常见的。

那么，当我们进入更深层次的物质结构，不再考虑整个原子，而将注意力集中到组成原子核的质子和中子的结构上时，为什么问题就变得极其复杂了呢？这是因为人们认为质子和中子是由夸克组成的，夸克在组成质子或中子时要保证总的色荷为零，这就像原子要保持电中性一样。但是从原子中取出一个或多个电子，从而形成一个离子是很简单的问题，而要在核子中进行类似的工作——从核子中拿掉一个夸克，留下一个“色离子”——就不仅仅是困不困难的问题了，因为那是不可能的，或者说起码看起来不可能。

其原因当然是因为夸克禁闭。夸克禁闭假定，只有颜色中性（或者称为“白色”）的粒子才能够在自然中观测到。因此仅从质子中取走一个夸克是不允许的；同样，一个自由运动的裸露夸克

也不允许存在。但这只是一个假设，该假设是为了解释没有观测到的自由夸克和胶子而提出来的。这个假定从来没有被完整地证明过。如果这一假设成立，强相互作用理论的主要成员将永远都不会被看到，这一假设一直得不到证明是一件令人很不满意的事情。禁闭就像是总在一间封闭小屋内开会的委员会，人们永远看不到也听不到它的成员。禁闭问题被认为是科学中最重要的一个尚未解决的挑战。

除了需要证明理论的这个核心假设之外，还有一些额外的因素。QCD 理论看上去对一些特殊的情形，例如质子之间发生的某些剧烈碰撞以及产生喷注的过程，解释得不错，但这些仅仅是一部分问题。最原始的问题很容易被人们遗忘，那就是在最基本的层次上理解物质。QCD 理论就是有关强相互作用的理论，它能够解释人们无法看到的基本实体——夸克之间通过胶子传递的相互作用力。实际上，尽管理论物理学家们要完整地解释强相互作用，但他们在考虑这项工作之前，必须能够利用 QCD 理论解释夸克和胶子是怎样结合成质子和中子的，因为质子和中子组成了原子核，而原子又构筑了我们的物质世界。他们还要能够通过理论预言核子的性质以及其他一些可观测粒子的性质。最终，这幅完整的图像反过来又要能够解释质子和中子怎样黏结到一起，形成了位于原子核心的原子核。

现在问题出现了，由于可观测粒子内部的夸克——例如核子、介子以及其他任何强子内部的夸克——都处于相对低能的环境中，它们之间的耦合作用是强相互作用，在低能状态下 QCD 微扰理论是不适用的。因此，解释相互作用着的夸克怎样结合成真实粒子以及证明夸克紧闭都是非微扰问题，其中“非微扰”可以翻译为“能量低，距离相对较长——大约为 1 飞米”。现在，在这种距离和能量上处理强相互作用要用到计算机模拟（这是一种称作格点 QCD 的领域，我们一直要到第十章才能介绍这方面

的内容），设计出一套灵巧的模型，用一些取巧的方式从不可能处理的问题中分离出能处理的一小部分问题。

夸克禁闭

一种简单的方法能够让我们理解粒子怎样存在、它们之间有什么关系以及一些关于夸克禁闭的提示，这种方法就是将一些QCD理论的内容放入夸克模型，并且以粒子分类图作为指导。

引起夸克碎裂的机制，有点像一幅简单的夸克禁闭连环画，它是这样实现的。想象一下，组成π子的夸克和反夸克，是由某种弦连接在一起的。夸克和反夸克中的一个会以很高的速度绕着另一个运动，它们之间通过弦保持距离不变。这是一种强度很大的弦，因为禁闭就表示在夸克和反夸克被拉开的时候，它们之间的作用力保持不变，当这种弦被拉长的时候，它不会变弱。加在两个夸克上用来使它们分开的能量，转变成了更多的弦，弦也会变得更长。

但是这种弦不能无限制地伸长。当能量达到某一个点时，弦会突然崩断，弦上新产生的一个端点会出现一个夸克，而另一个端点会出现反夸克——在第六章中，我们讨论夸克碎裂时已经谈到过这种图像。因此，现在出现了两个π子，每一个都是由夸克、反夸克和连接它们的弦组成的。对于重子来讲，一种更复杂的关于弦的连环画联系着三个夸克，但基本思想都是一样的，见图7.1。为了和QCD理论联系起来，人们把弦说成是由色力“纤维”编织而成的。

那么胶子是什么情况呢？在夸克和反夸克相距非常近的情况下，它们的行为就显得相对自由；基于重夸克及其反夸克的模型支持单胶子交换，并且满足同库仑定律一样的方式。但随着更多的能量注入各个夸克-反夸克对，这些能量会从真空中产生越来

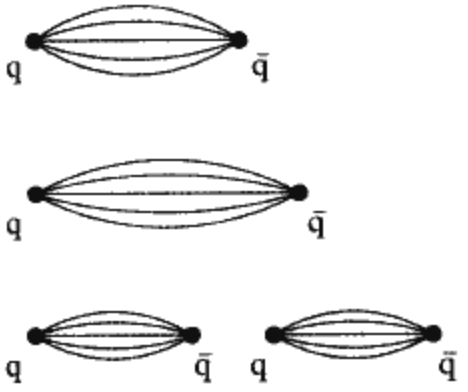


图 7.1　用弦描绘的夸克禁闭图。当人们把夸克和反夸克往两边拉的时候，最终系统能得到足够多的能量，从真空中产生另一个夸克-反夸克对

越多的粒子，也包括胶子。在相互吸引的夸克和反夸克之间有传递相互作用的胶子，它们之间存在着一种吸引力，这是由大量云集的胶子造成的，等效于将所有胶子的色力线都吸收到一起。这是胶子本身带有色荷的直接结果，因此胶子之间才能够发生相互作用。越用力扯夸克-反夸克对，就会有越多的胶子参与到这种混战中来，也就会有越多连结夸克和反夸克的相互作用力线聚集到一起。夸克和反夸克之间的相互作用就不再是那么微弱，这些力都沿着一个作用力管的方向，从而产生了连结夸克和反夸克的“弦”。我们在第五章讨论过反屏蔽效应，它能够使夸克和反夸克所带的色荷看起来更强。除此之外，我们还引入了新的内容，那就是色荷力线能够集中到一个“管道”中。当夸克和反夸克比较接近的时候，反屏蔽效应更显著，随着它们之间的距离越大，反屏蔽效应使相互作用强度更强。当把它们之间的距离拉得更远一些，色力就开始被导入连结夸克与反夸克的管道，其作用强度随着距离的增加大致保持为常数，它最终导致了夸克对的禁闭。

对一些人来讲，这种图像已经足够了。但对于另一些人，还存在一套更基本的禁闭弦理论，许多人认为，这种理论能够为我

们更加精确地理解禁闭过程。该模型来自于凝聚态物理中的超导现象，许多物质冷却到特定的温度之下后，其电阻就会消失，这就是超导现象。在超导现象中出现的新内容就是一种由弱耦合电子对组成的“流体”，这有点像夏季的傍晚草地上凝结的露珠，高温下遍及整个金属的电子“草地”中也会凝结出电子对。这种流体（或者称凝聚态，就像降低某些气体的能量时发生的那样）中的电子对是一种密度更高的态，它通常被称作库珀对，是根据利昂·库珀（Leon Cooper）的名字命名的。他是赢得 1972 年诺贝尔奖的三位美国物理学家之一，他们三人因为解释超导现象而获此殊荣。库珀对的基本思想是，一个电子使它所在的晶格发生变形，另一个电子感觉到了这种畸变，并做出相应的调整以降低其能量。因此一对通过格点发生间接相互作用的电子，形成了一个总自旋为 0 的物体——一个玻色子，其所带的电荷是电子电荷的 2 倍。

库珀对是一种十分奇怪的东西。库珀对中的两个电子可能相距上百个原子间距，因此有大量的库珀对重叠。库珀对是玻色子的事实说明，它们不会受到约束电子的泡利不相容原理的作用，这就表示它们可以统一行动——它们就像在第三章提到的超冷却铷那样，会产生玻色-爱因斯坦凝聚。正是这种集体行为造成了电阻的消失，提供了一幅与传统电传导截然相反的戏剧性图像。在传统的传导过程中，金属内部电子气中的每个电子，都奋力地为自己在金属中开路，每走一步都会受到带电离子的撞击，就像一个想要去地铁站的匆匆过客在一条拥挤的大街上穿行。与之相反的是，在超导过程中的传导更像一支步伐整齐的检阅部队。

超导体具有各种令人吃惊的性质。其中与夸克禁闭的故事有关的，是一种在 1933 年被沃泽·迈斯纳（Walther Meissner）发现的一种效应。迈斯纳发现位于磁场中的超导体会阻止磁场通过它的内部。这种所谓的迈斯纳效应，是由于超导体表面存在电流

回路引起的，这些电流会在超导体内部产生与外加磁场大小严格相等但方向相反的磁场。

超导体可以分成两类。像水银这种金属称作Ⅰ型超导体，因为它能产生经典的迈斯纳效应，它能够将磁场完全排除在体外。而另一些材料，像铌等，被称作Ⅱ型超导体。这些材料允许在整块材料内部的许多非超导区域内存在一丝丝很细的磁场。每条这种磁力线周围环绕着电流回路漩涡，它们能够在材料内部产生量子化的磁场。因此Ⅱ型超导体只能部分地呈现出迈斯纳效应，材料的绝大多数部分都将磁场排除在外，磁场只能存在于一些狭窄的管道之内，周围的超导材料只有在这些地方才能够彬彬有礼地款待磁场。这就是场的丝状模型，场被周围的环境挤压到管道中，这种现象引起了研究 QCD 理论的理论物理学家们的兴趣。

要发展一套同Ⅱ型超导体理论相类似的 QCD 理论，还需要加上另一种要素。对于一个介子来讲，连接夸克和反夸克的弦会从这对粒子中的一个开始，在另一个粒子所在的地方结束。如果将束缚磁场的管道概念应用到这种介子弦模型中，在磁场管的一端必须存在源头，而磁场管的另一端也必须有一个“排水口”（称为“汇”）。换句话说，应该存在一个单独的磁极，比如在磁场管的一端是磁北极，磁场管的另一端就是与之相反的磁南极。现在问题出现了，因为尽管单个磁极（或者称磁单极，monopole）能够使麦克斯韦方程组更加优美，大自然看起来还是没有制造出存在单个磁极的机会。大自然要求，真实的磁铁——无论是条形磁铁还是螺线管——都必须拥有两个磁极：磁南极和磁北极。与这些具有两个磁极的磁铁（或者称磁偶极）相反，电场的起源是一种单极。也就是说，一个电场的源——正电荷可以单独存在，而电场的汇——负电荷也同样是这样。

不管怎样，发明出一种磁单极并假设它们存在于超导体当中，是一种十分有效的构建模型的方法。这样，磁单极就可以作

为磁场管的源头，与之对应的成员——反磁单极也可以作为磁场管结束端的汇了。这种磁场管的能量会随着其长度的增加而增加，磁场被束缚在管道中并且是量子化的，所有这一切都是禁闭模型中所需要的性质。

尽管这种模型具有这么多令人振奋的性质，但是它作为夸克禁闭模型来讲，仍然不能完全满足要求。原因在于在真实世界中，可以说是先有电荷，再有磁场的。也就是说，电子和正电子之间的相互作用力，首先要依赖于它们所带的电荷。磁场力是电荷运动的结果，从这种意义上来说磁效应是次要的。如果在 QCD 理论中，电荷所对应的是色荷，夸克和胶子正是通过色荷才能互相感知到对方的存在。

要进一步发展超导模型，所用的方法就是将电场和磁场交换。这种交换就是被数学家们称为“对偶交换”的一个例子，其结果被称为双超导：一个单极形成了电场的源，而一个反单极形成了接收电场的汇，周围的磁单极海洋维持着管道中的电场。作为一种 QCD 理论的夸克禁闭模型，单极被替换成了夸克和反夸克，它们分别成为了色电场的源和汇，色电场管道将它们联系起来，周围的色磁单极将场挤压到了管道中。这就是夸克禁闭的双超导模型。

这一模型的真实含义是，禁闭系统禁止夸克和反夸克成为自由粒子，是由于周围存在色磁单极。从色电单极凝聚的意义上来讲，介子和重子周围的真空是一种超导物质，色电场由于迈斯纳效应从周围的单极海洋中被排挤出来，挤进了连接夸克和反夸克对的狭窄管道。

20 世纪 70 年代的又一产物——双超导模型要归功于许多人的不懈努力，尤其是哥本哈根的霍格·尼尔森（Holger Nielsen）和保罗·奥尔森（Poul Olesen），CERN 的布鲁诺·朱米诺（Bruno Zumino），加利福尼亚的斯坦利·曼德斯塔姆（Stanley

Mandelstam)，莫斯科的亚历山大·波利亚科夫（Alexander Polyakov)，芝加哥的南部阳一郎以及乌德勒支的杰拉德·特霍夫特。该理论在理论物理学家中间仍然十分流行，他们知道至少从理论上讲，尽管实验物理学家们从来也没有发现过一个色磁单极，但如果它们存在，QCD 理论将会十分舒服。但夸克禁闭真的是由于充满整个宇宙的单极子引起的吗？没有人能回答。

强子模型

最简单的同时也是人们研究最透彻的原子就是氢原子，它有一个只包含一个质子的原子核和一个轨道电子。这个电子可以处于原子中许多不同的能级上，量子力学可以全面地解释这些能级的形成。用某些方法来激发原子，电子就会爬升到能级的不同阶梯上去。一段时间之后电子会跌下来，原子会发出光以及其他辐射，这些分立的频率反映了原子的能级。研究光以及其他辐射发散（或吸收）的分支称作光谱学，这是在一些未知样品中验证原子种类的有效方法，光谱能够提供一种“光指纹”。

可以设想一种“原子”，其中质子被一种具有与电子同样质量但带有正电荷的物体所取代，也就是将质子换成一个正电子。这看起来不太可能，因为与夸克和反夸克相反，电子和反电子常常发生湮灭，但这种“原子”是可行的。它被称作电子偶素。

电子偶素分为两类。一种寿命较长，称作正电子偶素（orthopositronium)，其中电子和正电子的自旋都指向同一方向。另一种是超电子偶素（parapositronium)，它是由自旋相反的电子和正电子组成的。尽管电子偶素的寿命特别短（超电子偶素的寿命大约为 1.25×10^{-10} 秒，而自旋相同的那种电子偶素寿命大约是它的 1000 倍)，但我们还是能够精确测量出它的能级光谱。电子偶素的能级同氢原子的极其类似，但能级之间的间隙比氢原子

小一半，这是由于原来质量很大的质子被质量小的正电子所取代造成的。

如果用更大的分辨率，氢原子主能级实际上是由许多具有细微差别的能级线组成的。有一些能级线可能是由于电子自旋引起的。电子偶素同样也具有额外的“精细结构”，因为电子“磁铁”会发现自己处于一个磁场中，这种磁场是由电子和正电子之间的相对运动所引起的。但我们还需要考虑到更多的谱线。这种“超精细结构”也出现在原子光谱之中，这是因为电子和正电子都具有自旋，因此它们的行为都像一个小磁铁，它们相互之间的磁相互作用依赖于它们各自的相对指向。超电子偶素的自旋相反，因此它的能量也较低，正电子偶素的能量仅仅比它高 1/1000 电子伏。电子偶素光谱的测量结果同计算结果十分吻合，电子偶素也是一种物理学家们理解得十分透彻的对象。

之所以要说到电子偶素，是因为它有可能为我们提供一种介子模型。理解介子光谱的一种方案是利用成熟的 QCD 计算方法，但是没有人知道怎样去做。一种经过简化但是仍富有成效的替代方案就是用电子偶素做模板：将电子偶素中的电子替换为夸克，正电子换成反夸克，基于光子的相互作用换成有关胶子的相互作用，就可以看出最终是否能得到一个介子。那么，问题就在于，介子质量能否表现出类似于电子偶素那样的能级样式，并且用于电子偶素中的简单量子力学有多少能够被粒子物理学家们用到介子物理中去？

其中会产生一些重要的差别。首先，因为介子中将夸克和反夸克结合到一起的力，比电子偶素中的电磁力要大得多，因此介子的能级间距相应地也要大得多。电子偶素中，电子和正电子所带电荷之间的静电力同强相互作用很相似，相当于两个静止色荷之间的“色电”力。因此电子偶素的库仑定律静电表述，可以用来描述夸克和反夸克所带色荷之间的类库仑定律色电力。此外，

夸克具有自旋，因此和电子一样也像一块小磁铁，能够发生电磁相互作用，但这种小小的贡献并不是最重要的。重要的是产生自旋的色荷也能够产生“色磁”力，这种色磁力同由自旋电荷产生的普通磁力相类似。这种色磁力在夸克和反夸克之间产生作用，它也依赖于自旋的相对方向，并且同电子偶素的能级阶梯相类似，表现为一系列的介子态。

观测这种色磁力效应的最好场所就是重子的质量模式。简单的夸克模型会留下一些尚未解决的问题。例如，早期夸克模型中经典的质量关系不能够解释，重子八重态和重子十重态中具有相同夸克结构的成员为什么质量会有差别。一个很好的例子就是质子，它的质量是 938.3 兆电子伏，夸克结构是 uud；Δ^+ 粒子的质量为 1231.6 兆电子伏，质子质量大约是它的 76%，但它同质子含有同样颜色的夸克。这种差别要归咎于夸克所带的色磁力。Δ^+ 粒子是重子十重态中的一个成员，组成它的三个夸克自旋指向都是相同的。而质子中的三个夸克，有两个自旋方向是相反的。两种粒子具有不同的色磁相互作用，因此导致了能量的不同，这就简洁而优美地解释了质子-Δ^+ 粒子之间的质量差别。另一个经典的例子是 π 介子，它的质量为 140 兆电子伏，由 u 夸克和 d 反夸克组成；它的质量要比质量为 770 兆电子伏的 ρ 介子小得多。这两种粒子同样也都是由具有相同味荷的夸克组成。其质量差别来自于色磁力的相互作用，π 介子中夸克和反夸克的自旋是反平行的，而在 ρ 介子中则是平行的。

将夸克塞到一个 π 介子大小的空间内，该空间可要比电子偶素的大小小得多，这样会导致额外的复杂性。根据不确定性原理，对于质量比较轻的夸克，夸克动量的不确定性相应地会比较大，这就说明夸克的速度接近相对论速度。π 介子以及其他由接近光速的夸克组成的重子都是相对论系统，这使得将它们内部的相互作用区分为自旋部分、与相对运动有关的部分及静电部分的

方法失效了。这种用于电子偶素的方法实际上忽略了狭义相对论效应，因此注定会失效。但是发生在 1974 年 11 月的革命改变了一切。大自然十分同情粒子物理学家们，为他们提供了 J/ψ 介子，它是由很重的粲夸克和反粲夸克组成的束缚态。因为粲夸克和它的反夸克质量特别大，所以它们离得特别近，因而它们之间的束缚会比较弱。与将它们黏结到一起的能量相比，它们的质量也显得十分巨大，因此它们运动得很慢，近似地将相对论效应忽略掉是合理的。此外，对于重夸克系统，我们有理由假设组成介子“海”的所有胶子、夸克和反夸克，对于那些运动迟缓的笨重（价）夸克来讲干起活来要麻利得多，“海”所施加的净作用力可以被抹平，从而得到一种束缚夸克和反夸克的平均相互作用。对于 ϒ 介子来讲，前景更加光明，因为它由质量更大的底夸克组成，相对论效应更不显著。

有一整类介子都很像电子偶素，J/ψ 介子就是其中的一员，它们被称为粲素态，或者简称为粲素（charmonium）。它们都是由同样的粲夸克和反粲夸克组成的，但是却具有不同的质量和不同的量子数。其中质量最小的是 η_c，质量值为 2979.8 兆电子伏，而 J/ψ 介子的质量为 3096.9 兆电子伏。η_c 介子中的夸克和反夸克自旋方向相反，而 J/ψ 介子中的自旋方向却是平行的，因此 η_c 介子和 J/ψ 介子之间的关系特别像超电子偶素和正电子偶素之间的关系。由于 η_c 介子中的自旋方向是相反的，它所具有的量子数与在电子-正电子对撞中直接得到的不同。例如，它能够在电磁衰变过程 $\psi(3686)\rightarrow\eta_c(2980)+\gamma$ 中出现。在 3730 兆电子伏的质量以下，会出现一系列寿命相对较长的 cc 态，包括 χ_{c0}、χ_{c1} 和 χ_{c2} 介子以及被证实的 $\eta_c(3594)$ 介子。3730 兆电子伏这个数值十分重要，因为在这个值之上的粲素质量太大了，它们都会衰变为带粲荷的 D 介子，其中最轻的是质量为 1864.6 兆电子伏的 D^0 介子。那些不能迅速衰变为 D 介子的比较轻的态都能够精确地测

量，它们的质量谱可以在表格中列出来，摆在电子偶素质量的旁边。建立在底夸克基础上的 ϒ 介子系统的低质量态也有同样的规律，这些态被称为底素。如果顶夸克的寿命足够长，基于顶夸克和反顶夸克的介子会表现得更出色——它会变得和底素一样。

在开始产生 D 介子和 B 介子之前，电子偶素可以为粲素和底素提供十分满意的原型。介子的质量谱可以用各种量子数来标记，虽然它与电子偶素的能级谱所涉及的实际能量有很大差别，但两者十分类似。粲素和底素系统的基本物理图像都是一样的，从粲素研究中得到的结果可以用来预言一些底素态。实际上，这种方法也有助于用夸克和反夸克来描述介子，并且有助于用具有自旋并带颜色的夸克和反夸克之间的色磁相互作用，来理解相应介子质量。

至少对于比较轻的态，使用满足单胶子交换的类库仑相互作用就足够了；但是质量更大的态就需要更多的内容。理论物理学家们采用逆向思维的办法，将观测到的粒子质量硬塞到模型中，他们发现了一种修正的相互作用能量，在处理质量大的粒子时效果更好。如果夸克和反夸克之间的距离加倍，这种相互作用能的额外部分也会加倍。这是一条很关键的线索。

能够实现夸克禁闭的一种方法，就是确保将夸克-反夸克对分离所需的能量是无穷大。而使该条件成立的一种方法，就是保证夸克之间的能量随着它们之间的距离同步增长。要将它们分开就需要注入特别多的能量，这些能量足以使系统能够“重新排列”。我们没有得到分裂出来的自由夸克和反夸克，而是像上面所讨论的那样，从真空中又拉出了另外一组夸克-反夸克对。粲素和底素表示在短距离时进行单胶子交换，当夸克、反夸克之间的距离较大时，粲素和底素会达到更高的激发态，相互作用强度增强从而引起夸克禁闭。

这种随着距离增高的能量——或者说是线性增加的能量，同

作用力保持不变完全一样。这就像有一条连接夸克和反夸克的绳子，当把这两个粒子往两边拉时，绳子上的张力保持不变。将模型应用到粲素和底素态上，我们可以推导出绳子中的张力大约有14吨重！这就像是一辆载满货物的卡车拉一条比质子尺寸还小的绳子一样。强力真的很强。

重子口袋

关于由重夸克组成的介子，我们就介绍这么多。那么普通重子，尤其是质子和中子又是什么情况呢？现在，解释它们仍是一种艰巨的任务，因为重子中存在三种主要的（轻）夸克，而不是像介子那样只存在两种，即使是“简单”的存在相互作用的运动三体问题，要处理起来也十分困难。我们缺乏一套能够直接解决重子问题的成熟数学方案。

我们可以利用一些技巧来尝试构建重子的QCD模型。最早将QCD理论应用到核子中的尝试是麻省理工学院的口袋模型，是由肯尼斯·约翰逊（Kenneth Johnson）和他的同事于1974年在麻省理工学院提出来的。他们的想法很简单：唯象地引入禁闭，把重子看做是口袋或者气泡，夸克不能从中逃逸出来。我们可以想象一个水壶，水壶中的水马上就要烧开了，从液态水中会产生蒸汽气泡，每一个气泡中的水分子之间存在较弱的相互作用。与之类似，口袋模型把重子或者介子看做真空中的一个气泡，这些重子泡中是自由运动的夸克。但是要产生一个强子泡，要比在茶壶中产生一个气泡需要更激烈的作用。例如，使用粒子加速器就是一种将足够的能量注入足够小的空间内，从而吹出夸克泡的方法。产生夸克泡所需要的能量称作口袋常数，每立方飞米夸克泡的体积内大约需要55兆电子伏的能量。至关重要的是，这种填充能量是和夸克泡的体积成比例的，因此要想继续使夸克泡膨胀，就要持续不断地向它输送无穷多的能量。夸克泡之所以

不发生爆炸或者释放能量，是由于夸克泡中的夸克快速地到处运动，因此在内部产生很大的压强。颜色禁闭在总体上限制了色荷源——换句话说，也就是夸克和反夸克——和色力线永远不会穿出夸克泡表面，它们总是被包含在内部。夸克泡总体上当然是色中性的。从数学上讲，该问题可以归结为解无质量夸克的狄拉克方程，其中要包括夸克泡表面的特殊情况。

“口袋工人”们已经获得了成功，他们成功地解释了强子质量和总自旋之间的关系以及强子质量的基本图谱。他们还能够给像介子碎裂这样的过程提供一套独特的模型。但这次他们的成功也并不是独一无二的——上面谈到过的弦模型也瓜分了一些他们的地盘。弦和口袋看上去都有些粗糙，不太像物理上的内容，并且面临着许多困难，但是它们都具有一个优点：它们都能够抓住混乱问题的主线，理论物理学家们利用它们能解决实际问题，这对于理论物理学家们来讲是一件好事。

企鹅图和褐色的垃圾

用质子的各个组成部分来描绘质子的运转情况很难，这是因为根据 QCD 理论的标准，那些组分在一个相对很大的盒子里进行活动。对于紧紧拥抱在一起的夸克和胶子来讲，这是一个比较优雅的环境——在这里，纯微扰论对 QCD 理论是不适用的。

问题在于，是否存在一种环境，微扰论至少能部分适用呢？有没有可能从难以驾驭的长程、低能部分分离出短程、高能的部分呢？这样理论物理学家们就可使用他们强大的费曼图微扰理论了。答案是肯定的，能够处理这方面问题的数学工具被称作“算符乘积展开”（operator product expansion），或者简称为 OPE。

OPE 思想把物理现象小心地根据长度分隔开。用长度尺度来刻画系统的性质，比指定某种“背景”过程的截断假设要好得

多。短于截断长度范围内的性质对“相似性测量”（likelihood measure）有贡献，这种测量决定了大尺度效应的相对重要性。

我们可以举一个包含两种不同尺度过程的有趣例子，那就是电子对质子的深度非弹性散射过程。其中，被交换的探测光子波长同质子的整体大小相比显得特别小。因此，光子波长和质子大小是两种迥然不同的长度尺度。

如果给一个特写，深度非弹性散射就会以另一种样子出现在我们面前，那些散射组分通过一种神秘的方式结合在一起形成一个质子，可以用相似因子（likelihood factor）描述质子内部散射组分的分布。相似因子依赖于探测光子的波长，它等价于描述夸克在质子内部分布的结构函数。散射中心在一种神秘的环境下以某种方式组成了质子，我们要想了解全部内容，就需要一套可以解释夸克是怎样结合成质子的完整低能理论，但我们现在还对此一无所知。不管怎样，把质子分割成散射中心的方法仍然是有效的，那些结构函数我们在第八章还要提及，尽管我们不能够从头开始把它们计算出来，它们仍揭示出大量有关质子的信息。

尽管乍一看上去，OPE 像是一种简单的分割技巧，但实际上远不止如此。我们能够看出它包含了哪些额外内容，这种划分是基于尺度的，划分所选择尺度的长短具有任意性，但反映的物理内容却是一样的。这就是重正化的领地。实际上，相似性测量受重正化群的支配。这就表示，一旦在某个难以察觉的精细尺度上可以用微扰论计算相似性，我们可以将摄像机“拉远”，从而得到尺度膨胀时的情形，这也许能够符合实验观测。

我们可以用战争做类比。假设作战双方的军队都按传统的方式进行编制，我们可以利用一种中等尺度的单位来衡量，一个排假如由 30 个人组成，它包含 3 个班，每个班有 10 个人。而反过来，每个排又都是由中尉领导的，这些排是更大的作战单位——连的一部分，几个连又能够组成一个营。这样，中尉就可以从他

"排长尺度"的信息来理解他自己的战斗情况，从而估计国家尺度上更大问题的相对重要性，例如交战双方的技术综合水平、国内的公众情绪以及是否得到同盟国家的支持等。作为一个特例，如果一名中尉自己的部队装配着 M16，而敌军却拿着打猎用的老式来福枪，他就可以当场看到自己部队的作战效果。

用排作单位来分析战争是一个比较小的尺度——但我们还可以用更小的单位，比如说班，或者单个士兵——用光子来分析质子也是一个很短的尺度。战争的背景——例如技术水平、公众情绪以及同盟国支持等——可以比作由一个完整质子所提供的散射背景。我们甚至还能够想象出一个部队中的重正化群的样子，它表示怎样将一个排的经验提升到一个营的高度，并对战争的前景向将军或者新闻媒体说出自己的观点。

算符乘积展开实际上是在 QCD 理论之前提出来的。1969 年肯尼思·威尔逊（Kenneth Wilson）将其引入了 QCD 理论，具有讽刺意味的是，肯尼斯本来的目的是要用它代替场论。它的第一个重要结论是哥伦比亚大学的诺曼·克里斯特（Norman Christ）和他的同事在 1972 年得到的；两年之后，格罗斯和维尔茨克利用 OPE 和重正化群，讨论了深度非弹性散射过程的比约肯尺度不变性和（很重要的）对完美尺度变换的偏离，这个问题我们将在第八章仔细讨论。

苏联革命

如果我们近距离观察光子扰动真空的情况，就可以看到 OPE 怎样逐步成为了一套能计算强相互作用粒子性质的强大工具。例如电子能够和另一个电子发生散射，这是一个特别简单的电磁过程，可以用交换单个虚光子的费曼图来描绘。通过量子起伏，虚光子可以以夸克-反夸克对的形式度过其生命的一部分。这就是在第三章中讨论过的真空极化：图 3.8 中的图（b）是经典的真

空极化费曼图，它能够为 $e^- + e^- \rightarrow e^- + e^-$ 过程散射截面的简单 QED 计算提供修正。

当光子的波长特别短时（这十分接近于真实的情况），夸克-反夸克对将在真空中产生，过了很短的时间和距离之后，它们又被真空吸收回去。但是如果量子诞生和灭亡之间的间隔再稍微长一些，会发生什么事情呢？这就是在电子和正电子相互摧毁、发生湮灭的过程中，与虚光子相关的夸克-反夸克对所经历的事情。随着整体尺度的增加，在短距离上被渐近自由所压制的夸克-胶子相互作用变得越来越重要。这时，夸克和反夸克能够变成真正的物质粒子，例如 ρ 介子或 ω 介子，甚至是由重夸克组成的 J/ψ 介子或 Υ 介子，它们表现为电子-正电子湮灭截面中出现的共振峰。

理论物理学家们所面临的问题，就是怎样从我们已经透彻理解的过程外推，将直接观测到的量子起伏及其夸克圈图，外推到更大的距离上发生强相互作用的真实粒子。这就是三位苏联理论物理学家对自己提出的问题，他们分别是莫斯科理论实验物理研究所（ITEP）的米克希尔·施夫曼（Mikhail Shifman）和瓦雷丁·扎哈罗夫（Valetin Zakharov）以及位于诺沃西比尔斯克的巴德克尔（Budker）核物理研究所的阿卡迪·维恩施坦恩（Arkady Vainshtein）。他们认识到，这个问题的答案就隐藏在 OPE 中。

当夸克和反夸克的波长很短的时候，用微扰论处理问题是很安全的。当它们的波长增长时，麻烦就出现了。这时，夸克-反夸克对陷入了一层胶子和其他夸克与反夸克的包围当中，这些粒子都是从真空中来的。这就是根据尺度进行划分的开始，是 OPE 的核心。施夫曼和他的同事认为，在比较长的尺度上，这些夸克和胶子的效应能够累加到一起，（近似地）看作一个常数，作为系统的背景，但该常数尚无法计算。短程的夸克-反夸克对类似于一种传感器，对真空中距离相对较长的背景夸克和胶子有反

应。在短距离上，普通的微扰论就能描述夸克和反夸克，并可以通过 OPE 计算出相似因子。

从本质上讲，一个强子在短距离上看上去，就是一个逐步包含越来越多（非微扰）背景的序列。实际上，它并不是一个单独的强子，而是一整套量子态，最开始是像 ρ 粒子和 J/ψ 介子这样的基态，接下来是一系列与之基于同样夸克-反夸克组合的激发态及其他物质——毕竟没有人知道，从真空中产生的夸克和反夸克究竟打算形成哪种粒子。结果累加起来就形成了对强子参数的限制。换句话说，它是一种求和规律，是一种 QCD 求和规律方法，这种方法最初是由施夫曼和他的同事在 1978 年提出来的(图 7.2)。

图 7.2 米克希尔·施夫曼（左）、瓦雷丁·扎哈罗夫（中）和阿卡迪·维恩施坦恩（右）在他们研究 QCD 求和规律的期间摄于苏联。图片来源：米克希尔·施夫曼

因此，QCD 求和规则方法的出发点是 QCD 理论能够适用的地方，因为渐近自由在那里成立。结果显示，当把 QCD 理论推进到夸克-胶子相互作用更强的地方，求和规则的根基开始动摇，基于真空的非微扰贡献开始系统地出现。

把一个难处理的问题转化为另一个难处理的问题，而依靠真

空魔术般地凭空变出一个答案，从而推动理论继续前进，这可能吗？实际上，这是可能的。马上，我们将会在第九章中发现，真空实际上不仅仅是一个空荡荡的空间。其次，尽管真空中夸克和胶子的平均作用并不能在冷环境下进行计算，我们可以从一些实验过程中——例如 π 介子衰变过程中——得到有用的数值，测量过程通过某种方式将衰变和真空性质联系到了一起。举一个例子，QCD 求和方法能够算出 ρ 介子的质量为大约 750 兆电子伏，对比实验测量值为 770 兆电子伏；理论算出的衰变常数值为 213 ±20 兆电子伏，而由衰变过程 $\rho \to e^+ + e^-$ 确定的实验值为 216±5 兆电子伏。实际上 QCD 求和规则方法为理论物理学家们提供了能计算许多性质的强有力工具，例如可以用它计算夸克的质量、介子和重子的衰变性质、磁矩以及其他许多性质。施夫曼和他的同事最初发表的论文在 1979 年的《核物理》（*Nuclear Physics*）B 刊中占了整整一卷，现在它已经在科学文献中被引用了两千多次，这使得它成为最经典的 QCD 理论论文之一。

施夫曼和他的同事在 20 世纪 70 年代的工作取得了丰硕的成果，这是他们在克服了许多困难的情况下取得的一次巨大成功。由于要接受审查，期刊从西方引入苏联要延误很长一段时间，得到在西方杂志上发表论文的许可也需要经过好几个月。根据施夫曼的叙述，提交发表结果前的版本——即预印本，同时向杂志提交论文是不允许的。当时苏联法律不允许一篇论文同时作为预印本和正式出版物出现。因此，苏联的作者们经常修改论文题目和作者名单，以使检察员认为预印本和发表的论文无关。提出 QCD 求和规则的那篇论文的预印本，实际上变成了一连几篇陆续发表的系列文章，这样做是为了绕过文章长度限制审查。

除了以上这些困难之外，他们当然还面临着一些常见问题，这是每一个向期刊投稿的作者都要面对的，那就是评审过程。尽管施夫曼和他的同事最终在《核物理》B 刊上发表了他们求和规

则的工作，由于他们之前投的一篇论文被杂志和审稿人冷落了两年，他们曾经打算换掉那篇论文。先投稿的那篇文章中提到了企鹅图。

强力世界的弱力窗口

在奇异 K 介子衰变的奇异世界中，一切都不那么顺利。尤其是在 $K^+ \to \pi^+ + \pi^0$ 过程中，K^+ 介子能够衰变为一对 π 子。过程中没有观测到轻子，最终只存在发生强相互作用的粒子，因此该过程被称为非轻子衰变。中性 K 介子也能够衰变为两个 π 子，即 $K_s \to \pi^+ + \pi^-$。中性 K 介子具有一种奇怪的行为，这就是要标注符号 S 的原因，它代表“寿命很短”——中性 K 介子的寿命很短，大约是长寿命 K 介子或者称“L 型”K 介子寿命的 1/1000。长寿命的 K 介子一般会衰变为 3 个 π 子。

但是我们不用担心三 π 子的情形，只有二 π 子的情况才更有意义。实验物理学家们了解到，衰变过程 $K_s \to \pi^+ + \pi^-$ 发生的可能性要比衰变过程 $K^+ \to \pi^+ + \pi^0$ 高大约 450 倍。这种神秘的现象之所以有趣，是因为这两种情况都涉及弱衰变过程，并且组成粒子的夸克几乎完全一样，它们的性质也本应该大致相同。实际上早在 1964 年，朱利安·施温格（Julian Schwinger）估计出现 K_s 衰变的情况应该是 K^+ 的 9/4 倍。450 和 9/4，这两个数字之间的差别也太大了，因此它拥有一个著名的名字。但在说出它的名字之前，我们还需要先介绍一些相关内容：在 $K^+ \to \pi^+ + \pi^0$ 的情形中，总同位旋改变了 3/2，而在 $K_s \to \pi^+ + \pi^-$ 的过程中，总同位旋只改变了 1/2。因此“同位旋改变 1/2 的情况”可以记作 $\Delta I = 1/2$（其中 ΔI 部分是“同位旋 I 的改变”的缩写），它比 $\Delta I = 3/2$ 的情况发生的可能性要大得多。$\Delta I = 1/2$ 的路径，即 $K_s \to \pi^+ + \pi^-$ 过程占统治地位。这就是“$\Delta I = 1/2$ 定律”。怎样理解这种现象仍然是一个尚未解决的问题。

最重要的是，并不存在一种巧妙的对称性方法可以解释这个定律。解决问题的技巧就是承认强相互作用的地位。强相互作用效应通过两种方式起作用。首先，衰变依赖于 W 玻色子，由于 W 玻色子都很重，因此它们会进行短程相互作用，它们的交换力要受到短程胶子效应的支配。其次，在较长的距离尺度上，出现的夸克将会重新形成可观测的 π 子，强相互作用也会在这里出现。

但是如果在该问题上加上正的自旋，从另一个角度看这个问题——例如 K 介子的衰变过程——就会发现，弱相互作用实际上是一种获取强子内部强相互作用信息的方法。弱衰变为强子间的强力作用提供了一个窗口。

K 介子衰变是一个特别生动而重要的例子。为了在乍看上去是纯弱相互作用的过程中理解短距离强相互作用的影响，费米实验室的玛丽·杰拉德（Mary Gaillard）和本杰明·李（Benjamin Lee）以及罗马的圭多·阿尔塔莱利（Guido Altarelli）和卢奇亚诺·麦阿尼（Luciano Maiani）都试图用 OPE 寻找解答。这些先驱者的努力成果在 1974 年公布于世，他们开拓了一个新的领域。

W 粒子依靠它巨大的质量，为我们划定了“短程”的标准。

如果将这种短程尺度之下的细节——其中也包括几种重夸克——统统打包成各种各样的“常数”，最后只剩下三种轻夸克、光子和胶子，这样就利用在第三章中提到过的思想，将理论修改为一种有效场论。这看上去好像又退回到费米有关弱相互作用的旧理论，但实际上是一种推广。其中出现了一种神奇的新要素，那就是用重正化群武装过的 OPE。它说明了在保证隐含因素（例如 W 玻色子）影响的前提下，如何构造一套合适的有效理论。QCD 理论提供了一种在短距离上处理强相互作用贡献的有效描述。在短距离上，由于存在渐近自由，强耦合参数的值很小，可以把用 QCD 理论计算出的“常数”应用到较大的长度尺度上，

最终得到满足重正化群的有效理论。

杰拉德、阿尔塔莱利和他们各自的合作者发现，强相互作用确实能够加强 K_s 与 K^+ 衰变可能性的比率。这太棒了。但不幸的是，他们发现的加强程度远远不够。而在地球另一端的莫斯科和诺沃西比尔斯克，施夫曼、维恩斯坦恩和扎哈罗夫意识到他们遗失了某些东西，那就是其他理论物理学家都没有重视的、最近才发现的粲夸克的影响。把粲夸克包括进去后，他们在 K 介子衰变的有效理论中，得到了一种全新的图像。这就是企鹅图。图 7.3 中的图（a）所示的就是按照施夫曼和他的同事在 1975 年的提议画出的企鹅图，图中一个夸克圈通过一个胶子与另一个夸克圈相连。他们发现，新的企鹅图只对 $\Delta I=1/2$ 的路径起作用，这进一步解释了 K_s 衰变的增强效应。即使是这样，总的增强程度还是不够，$\Delta I=1/2$ 定律仍然是一个神秘的东西。但不管怎么说，人们现在拥有了企鹅图，它的真实性是通过 B 介子衰变被首次证实的，它对于大量的弱衰变过程具有十分重要的意义。企鹅图对我们理解 B 介子衰变过程的 CP 破缺至关重要，人们于 2001 年在 BaBar 探测器和 BELLE 探测器上首次观测到了这种现象。

但为什么把它叫做企鹅图呢？在 1977 年，CERN 的理论物理学家约翰·埃利斯（John Ellis）正在同玛丽·杰拉德、迪米特利·那挪保罗斯（Dimitri Nanopoulos）和塞吉·鲁达斯（Serge Rudaz）共同研究 B 介子的性质，那时基于 b 夸克的 Υ 粒子刚刚在费米实验室发现。他们提出了一套同那些苏联物理学家们相似的方法，相应地也提出了一套他们自己的费曼图，见图 7.3 中的图（b）。一天晚上，埃利斯、鲁达斯和 CERN 的学生梅丽莎·弗兰克林（Melissa Franklin）一起去酒吧。埃利斯和弗兰克林一起玩飞镖，他们打赌说，如果埃利斯输了，他就要把“企鹅”这个词放入他的下一篇研究论文中。弗兰克林中途有事必须离开，塞吉·鲁达斯自己完成了游戏，并且她打败了埃利斯。梅丽莎·弗

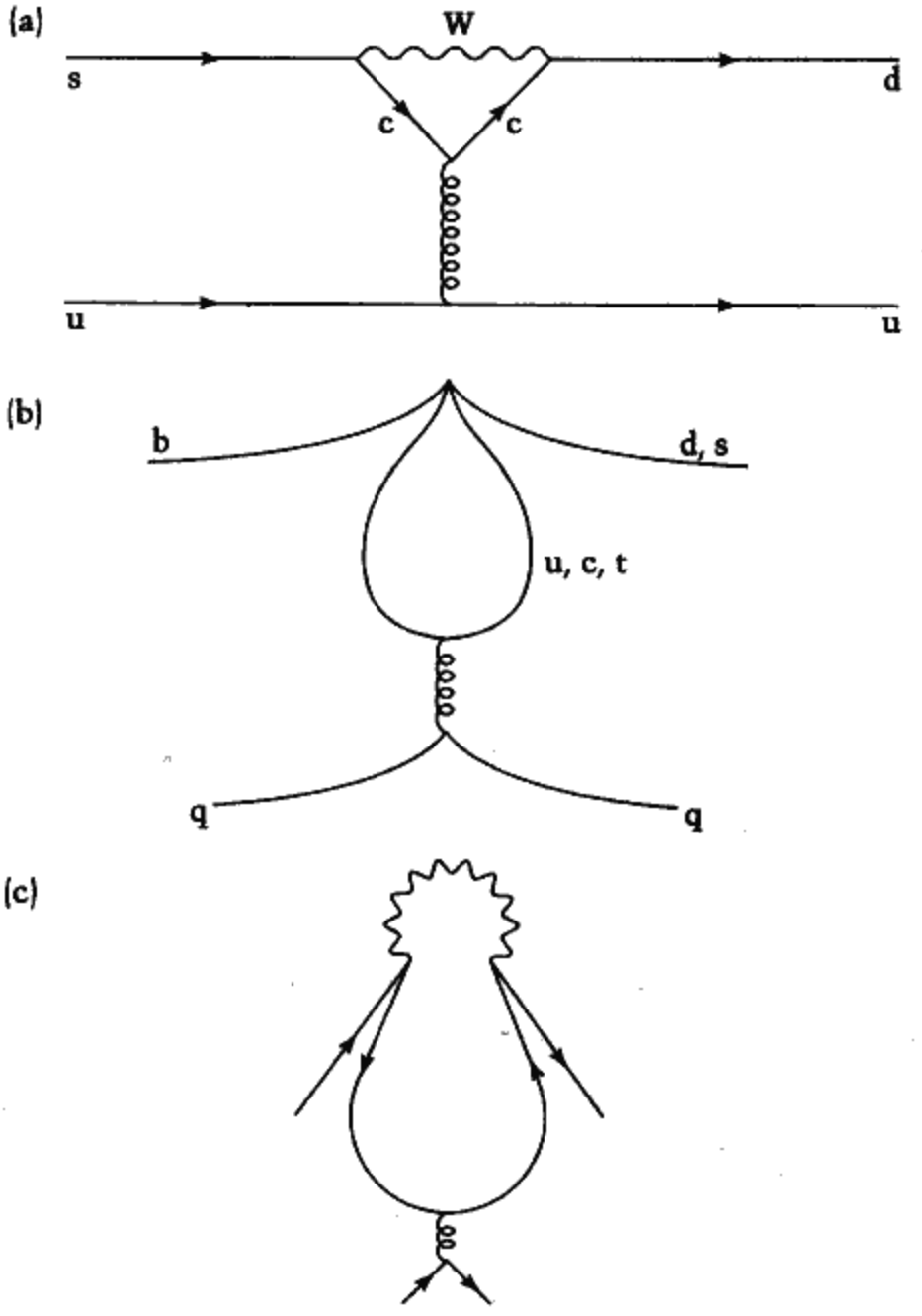


图 7.3 三幅企鹅图。图（a）显示了施夫曼和他的同事在研究 K 介子衰变时用到的一张企鹅图；图（b）是埃利斯和他的同事构想的一幅与 B 介子衰变相关的企鹅图。在图（a）中能够看到的 W 玻色子线封闭起来，在有效理论中产生一个点状顶点；图（c）是图（a）的变形，从而使它符合“企鹅图”的名字。图（c）是由乔汉（Johan Bijnens）和“世界科学”组织提供

兰克林不得不遵守他的诺言，但他一时间却不知道该怎么办。一天晚上他想出了办法。离开 CERN 之后，他顺便走访了一些朋友，并在他朋友家里吸了一些违禁烟草。那天晚上，他继续研究他的论文，一瞬间他突然发现费曼图看起来特别像企鹅，因此这个名字就被写入了论文中，一直沿用到现在。在瑞士，企鹅明显是很少见的。

心情沉重的介子

ϒ 粒子的发现，为理论物理学家们提供了一种可以把玩的真正沉重的夸克。例如在 B 介子中，底夸克（或者反底夸克）的质量是其他轻夸克（或者反轻夸克）质量的几百倍。在这种情况下，氢原子是一种比电子偶素还要好的模型。这是因为在氢原子中，质子要比电子重两千多倍，通过合理近似，我们可以把质子作为一个不运动的正电荷源，它能够将电子束缚在自己周围。与之相似的是，B 介子中的底夸克以及 D 介子中质量稍微小一些的粲夸克，都是静止的色荷源，它们能够将其他运动很快的轻夸克束缚在自己周围。理论物理学家们将这种介子称作是类氢介子。

因此，在最简单的层次上，重夸克的质量不再重要。可以肯定的是，B 介子和 D 介子的性质是相关的，它们虽然含有不同味的重夸克，但这些差别所造成的影响很小。这很像原子物理，不同的同位素都具有同样的化学性质，尽管它们的原子核具有不同的质量。重夸克的自旋也急流勇退，这在原子物理中也能看出来。改变氢原子中质子的自旋，对应着氢原子辐射频率谱的一个著名的特征，即所谓的 21 厘米线，这是一种只有典型电子跃迁能量百万分之一的跳变。与之类似的是，改变类氢介子中重夸克的自旋，可以通过色磁力表现出来，它的效应也是可以忽略的。这两种情况中，产生这种现象的原因都是因为较重粒子的（色）磁矩被它们巨大的质量所掩盖，因此改变自旋取向仅仅能引起能

量阶梯上一次很小的变化。

在20世纪80年代晚期，莫斯科的米克希尔·瓦罗金（Mikhail Voloshin）、施夫曼以及不久后多伦多大学的内森·伊斯格（Nathan Isgur）和加州理工学院的马克·怀斯（Mark Wise）都察觉到了一种新的对称性。他们看出，如果将重夸克的质量提升到无穷大，就能得到一种简化的QCD理论，这种理论对于重夸克的味和自旋的变化具有对称性。这是完整QCD理论所不具有的两种对称性。这种理解激发了一种方法，许多人对这种方法的发展作出了贡献，包括爱德华（Edward Shuryak）、波利茨以及其他一些人。1990年，霍华德·乔治亚（Howard Georgi）通过努力将其明确转变为一种有效的计算工具，称作“重夸克有效理论”（heavy quark effective theory），或者简称为HQET。重夸克有效理论是一种新的有效场论，它能够根据新对称性，描述重夸克和其轻夸克同伴之间的强相互作用。理论物理学家们运用这种方法，能够计算出一些过程的严格结果，例如包含无穷质量的重夸克的介子衰变过程。然后，他们就能够带入重夸克的质量，因为它们的质量也特别大，因此这种对称性也近似成立，通过对无穷大质量理论进行对称性破缺修正得到结果。重夸克有效理论是一种处理某些特别困难的问题的强有力技术，它可以被用来探索重夸克和轻夸克组成的夸克-反夸克对的性质。实际情况是，重夸克周围环绕着大量无法计算的夸克和胶子，它们都以极高的相对论速度①运动，伊斯格把它们称为“褐色的垃圾”。在技术文献中出现这样的表达方式，表现了怎样用多彩的手段刈除数学上的荆棘，也反映出真实问题本身是何等的困难。

在K介子衰变的情况中，W玻色子确定了自然的短程尺度。现在该轮到重夸克了。可以预言的是，在重夸克有效理论下面也

① 译者注：这里指速度接近光速，必须考虑相对论效应。

隐藏着算符乘积展开；并且就像它的名字所示，重夸克有效理论同第三章中讲过的内容一样，是另一种有效理论。

重夸克有效理论可以提供不同重夸克介子之间的关系，尤其是B介子和D介子家族成员质量之间的联系，还可以解释重介子的衰变。例如一个夸克经过弱相互作用转化为c夸克，其整体过程可以表现为$B \to D + l + \nu$，其中l代表像μ子或电子这样的轻子。可以肯定的是，这是一种弱衰变过程，但进行衰变的夸克被禁闭在介子中；这意味着，将B粒子衰变过程解读为b夸克到c夸克的转换时，强相互作用不能忽略。因此在重介子衰变过程中，粒子物理学家可以用重夸克有效理论来理清弱相互作用和强相互作用之间的纠缠。新建的B粒子工厂中的设备能够产生大量数据，将重夸克有效理论之类的技术应用到这些B介子数据中，能够得到巨大的收获。

第八章 显微镜下的夸克

拥有一套描述夸克间相互作用的理论，是令人十分愉快的事情。如果能够解释像 Υ 粒子这样的奇特粒子的性质，或者粒子之间的作用模式，也十分令人满意。但我们要解决的最基本问题是怎样理解原子核运作的原理，或者是另一个“修正过的基本问题”，即是什么控制了质子的行为。夸克模型和 QCD 理论都无法完整地向我们解释我们想要知道的东西，无法告诉我们组成物质的主要砖块——核子的所有性质。

深度非弹性散射是一种了解更多内容的方法，我们在第四章中已经介绍过，实验物理学家们可以在这种实验中利用粒子束“显微镜”，观测质子和中子内部的夸克和胶子；此外，它还是一种检验 QCD 理论正确性的重要方法。自从 20 世纪 60 年代末 SLAC 的实验物理学家们首次发现质子中的散射中心是点状实体时，深度非弹性散射就已经出现在实验议程中。随着实验数据不断更新，现在的实验物理学家们已能够区分不同颜色夸克的贡献，甚至还包括没有出现在核子中的粲夸克和奇异夸克的贡献。他们还能够观测到反夸克。他们不断锻炼技能，怀着坚定的意志，开始逐渐揭示胶子的内容。即使是核子内部自旋结构的秘密，也逐渐公之于世。

海夸克和价夸克

把光照射到质子和中子上，只要探测光的波长比核子的大小适当短一些，就能够揭示它们的内部结构。这种实验可以用结构函数来描述，结构函数在一定程度上是一种简单的唯象描述方式，它包含了靶粒子的所有结构。随着夸克概念的引入，表示夸克动量在核子中分布的结构函数逐渐声名显赫。

结构函数——或者说，夸克分布函数——对质子和中子内部已知结构进行了大量简化。它表示核子动量在夸克之间的分配情况，这同国家财富在民众之间的分配情况类似。计算每个人的财富占整个国家总财富的比例，然后将财富比例数作为横轴，而将具有特定财富比例的人数作为纵轴，就可以将财富分配情况在图中描绘出来。通过结构函数的数据来理解核子的情况，同根据财富分布图来理解一个国家的经济状况完全一样。

正如几乎没有人能占据国家总财富的绝大部分，但却有许多人拥有特别少的财富一样，几乎没有夸克会占据核子动量的绝大部分，但却有相当多的夸克动量值很小。图 8.1 大致表现了质子所包含的所有夸克（以及反夸克）的情况，这就是著名的 F_2，它是从深度非弹性散射中得到的两个结构函数中最重要的那一个。从最一开始，它就表现出，质子并不仅仅是一个由三个孤立夸克所组成的包。因为如果是那样，每个夸克的动量都应当是质子动量的 1/3，而不是一整套连续谱，图 8.1 就会演化为某种动量百分比值以 1/3 为中心的峰，见图 8.2。很明显，质子比简单的部分子模型复杂。

三个价夸克携带着质子的量子数，因此它们可以用重子数、电荷、同位旋等量来表明质子的身份，对应着简单夸克模型所要求的 udd 组成方式。因为价夸克能聚集在一起形成质子，它们必

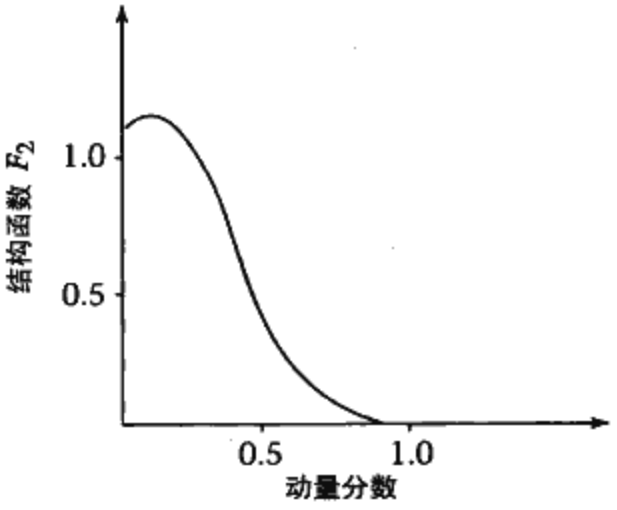


图 8.1 结构函数 F_2 的示意图，它是质子中包含的夸克和反夸克组分的动量构成

定要通过某种方式相互作用，很可能是通过胶子发生作用，这样才能让三个夸克共同分享质子动量。因此质子的动量不可能平分为 3 份，每个夸克占 1/3，质子动量会具有某种弥散的分布方式，图 8.2 形象地将这种分布表现了出来。起传递作用的胶子反过来也会产生夸克-反夸克对。这些粒子一般都会携带较小的质子动量百分比，它们会在夸克分布图中的低动量部分显著地出现。这是因为这些粒子所携带的只是胶子动量的一部分，而胶子只带有原始夸克动量的一部分，它们将质子所带的动量稀释了。正是这种描述为我们提供了一幅图像：价夸克沉浸在不断变化的低能胶子、夸克和反夸克“海”中。

有用的手性

如果这种描述是正确的，那么通过“海模型”，质子就会包含反夸克、奇异夸克以及粲夸克；而根据简单的部分子模型，这些东西在质子中都不存在。简单地用电子散射质子不足以揭示这些不同的可能性，因为它只能显示出夸克和反夸克的总情况。实

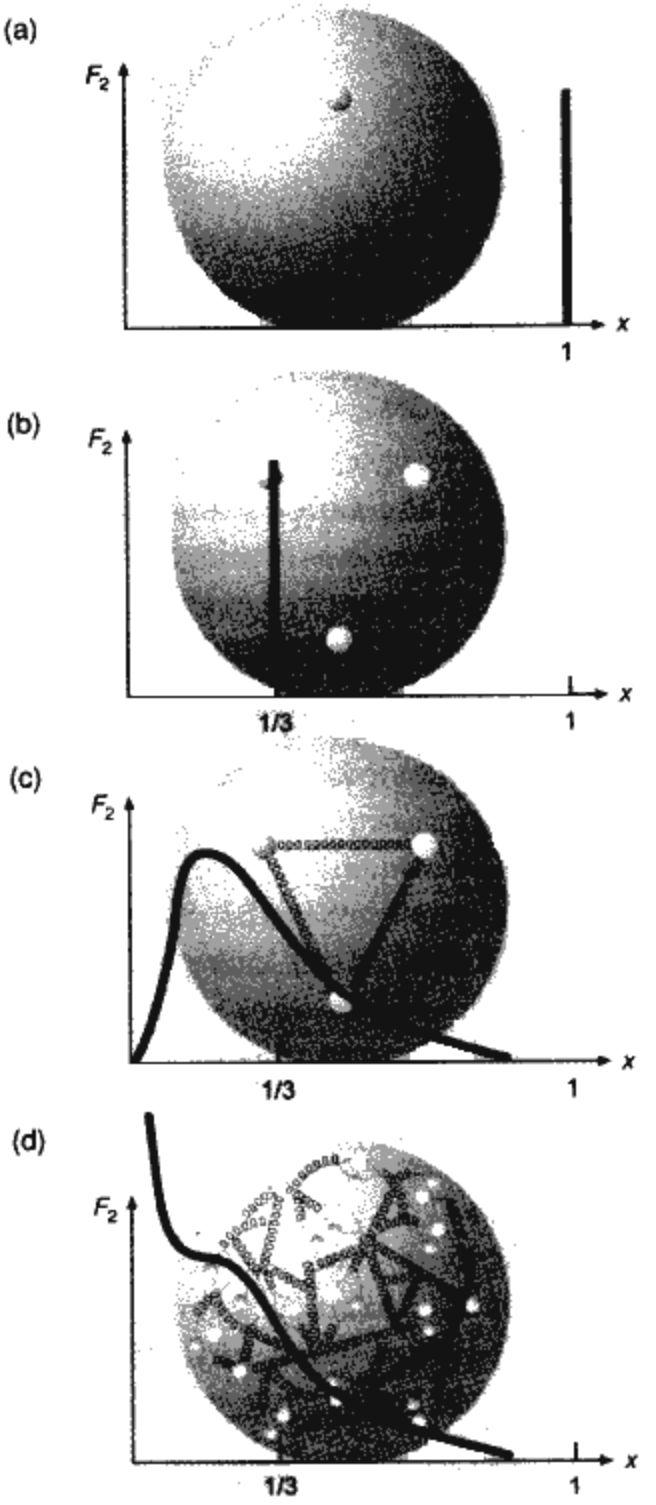


图 8.2 结构函数：如果质子由一个夸克组成，结构函数只能在 $x=1$ 处有一个值，其中 x 是夸克所携带的动量占质子动量的分数，见图（a）；而如果质子包含三个等价的自由夸克，结构函数将会在 $x=1/3$ 处显示出一个值，见图（b），因为质子动量在这三个夸克之间进行了平均分配；但如果三个夸克之间通过胶子以一种简单的方式相互作用，质子中的夸克分配就会发生改变，见图（c）；这些胶子反过来又可以生成海夸克和海反夸克，它们会在质子动量百分比很小的地方，引起结构函数激剧增大，见图（d）。图片来源：德国汉堡，DESY

验物理学家们怎样才能看到质子内部的奇特味道呢?

办法就在手边，中微子为他们提供了一种不可思议的方法。中微子的特殊性质之一就是不带电荷。在中微子深度非弹性散射实验中，入射的中微子会根据中微子类型相应地变为带负电的出射电子或 μ 子，见图 8.3。该过程通过刺入靶质子的媒介 W 玻色

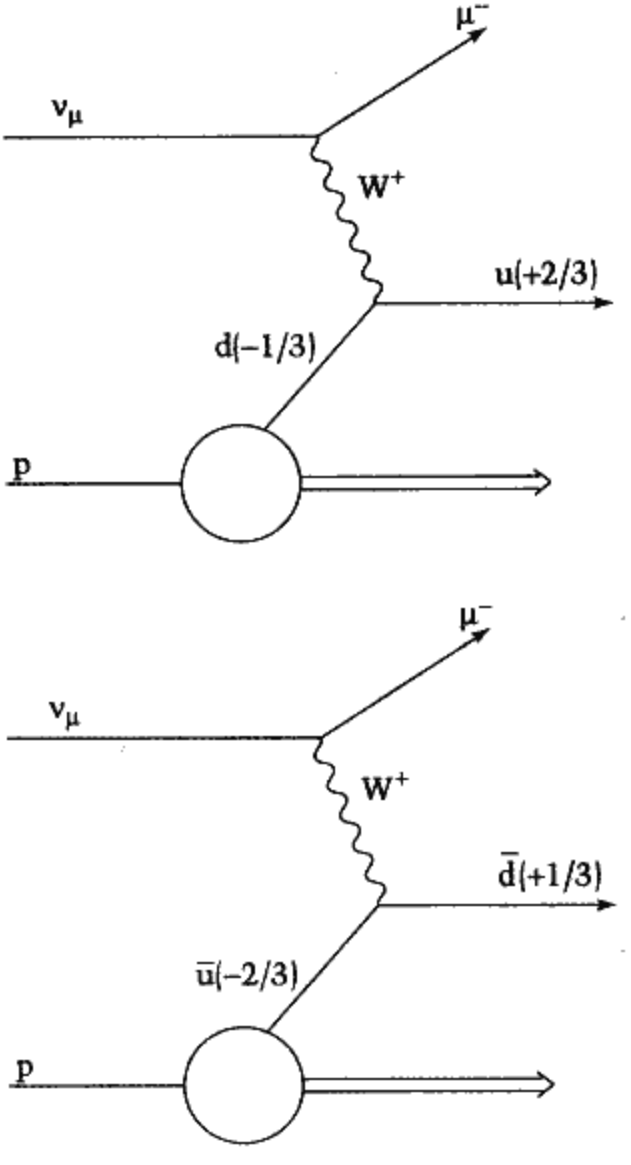


图 8.3 μ 子中微子同质子进行的深度非弹性散射过程。只有 d 价夸克或者 d，s，$\bar{u}$ 和 $\bar{c}$ 海夸克有贡献。夸克的电荷表示在圆括号中

子来维持电荷平衡。如果不考虑奇异夸克和粲夸克，电荷守恒说明只有下夸克和反上夸克参与了该过程。同样，对于反中微子散射，只涉及反下夸克和上夸克。

在研究核子结构时，令中微子如此特殊的第二个关键性质，就是它的手性（handedness）。左旋的中微子在大自然中占绝大多数，并且左旋的中微子只能同左旋的夸克和右旋的反夸克发生相互作用，这是弱相互作用“极力”破坏镜像对称性（宇称）的结果。第三，由于存在宇称破缺，结构函数必须包含手性。

手性以及通过带电玻色子探针作用的能力，使得中微子和反中微子散射对味道和正反夸克极为敏感，电子和 μ 子散射却对它们不敏感。这些性质帮助实验物理学家们理清了夸克同反夸克以及各种味道夸克的贡献问题。其中最为重要的是第三项镜像对称破缺结构函数，该函数本身可以通过中微子和反中微子对质子-中子混合靶散射概率的差别计算出来。根据简单的部分子模型，该结果应该等同于夸克同反夸克之间散射概率的差别。因为从“海”中出现的夸克应该同海中的反夸克对应，夸克海的贡献相互抵消，最终剩余的结果等于价夸克的贡献。这是求和规则的一个例子，即基于对夸克动量贡献求和的一种关系。这种独特的规则最初由普林斯顿的戴维·格罗斯和 CERN 的克里斯·卢埃林·史密斯（Chris Llewellyn Smith）在 1969 年提出，那是一个求和规则盛行而部分子思想尚处于酝酿当中的时期。格罗斯-卢埃林求和规则是这类规则中最精准的一个。例如，CCFR 小组（由芝加哥大学、哥伦比亚大学、费米实验室和罗彻斯特大学的科学家组成的合作小组）进行的中微子实验检验了该规则，在考虑了 QCD 修正之后，得到价夸克的数量为 3，误差在 10%的范围之内。因此，中微子散射证实，价夸克的数量符合简单的夸克模型。

实验物理学家们也可以比较中微子散射实验和电子散射实验

得出的夸克分布，从而检验夸克图像中另外一种基本成分。在差一个简单的数字因子的情况下，两个过程中的分布是一样的。这首先为整套方法提供了一个基本的检验，表明核子组分的分布不依赖于观测它们的粒子显微镜类型。其次，也是对简单夸克模型最为重要的一点，联系两者的数字因子等于具有预期分数电荷的夸克数。根据前两个结构函数之间的简单关系——即卡伦-格罗斯关系，深度非弹性散射为散射中心是自旋为 1/2 的点状粒子提供了证据，这些散射中心所带的电荷是 1/3 的整倍数。

伽格米勒气泡室小组在 1973 年首次得到了夸克海以正反夸克形式存在的证据，他们比较了中微子和反中微子对质子-中子混合靶的散射。中微子可以辨认“外来”夸克种类——例如奇异夸克，这为夸克海的存在提供了进一步的证据。中微子在同核子散射的过程中，可能会撞到一个奇异夸克，见图 8.4，假设它是一个 μ 子中微子，它就会变成一个 μ 子。同时，W 玻色子会同奇异夸克发生相互作用，产生一个粲夸克，粲夸克又会衰变成一个正 μ 子（也就是反 μ 子）、一个中微子和一个奇异夸克，这个奇

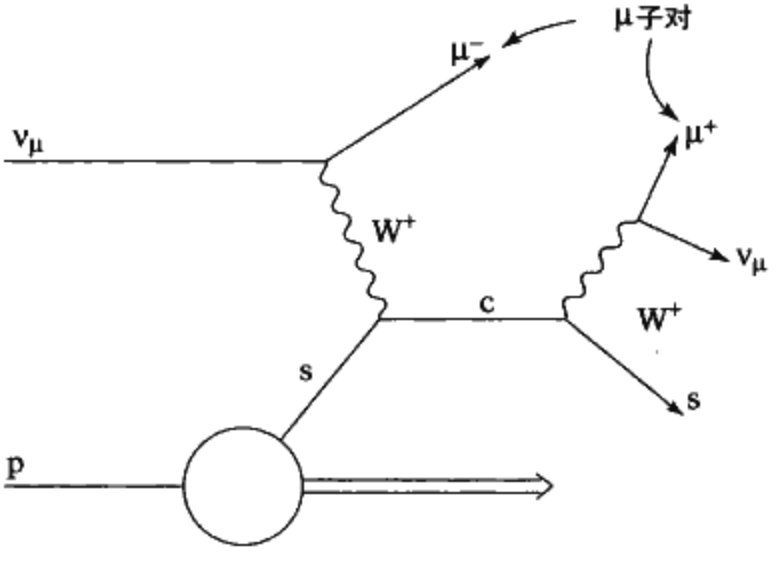


图 8.4 中微子与一个奇异海夸克发生深度非弹性散射，产生一对缪子（μ 子）

异夸克最终会产生一次强子喷注。该过程的标志，是同时产生两个符号相反的 μ 子——包括费米实验室 CCFR 合作小组在内的实验物理学家们都曾寻找过这种“双 μ 子”特征，并且都找到了。因为 d 夸克也会以类似的方式相互作用生成粲夸克，一旦 d 夸克的贡献被砍掉，就只剩下奇异夸克的信号。核子内部奇异夸克的分布，同其他较轻的海夸克类似，但奇异夸克的数量更少，人们认为奇异夸克和反奇异夸克的存在数量是相等的。

甚至还存在在核子中探测到粲夸克的可能性。一个 μ 子中微子同一个粲夸克通过中性的 Z 玻色子发生相互作用，会留下一个中微子，见图 8.5。粲夸克从容身之处被撞出来，衰变为一个奇异夸克、一个正 μ 子和一个 μ 子中微子。该过程的标志是一个“符号错误”的 μ 子。一个入射的中微子会导致出射一个正 μ 子，即反 μ 子。如果 μ 子仅仅通过交换 W 玻色子发生相互作用，轻子数守恒要求产生一个 μ 子，而不是反 μ 子。但是，要完成这个实验就意味着，既不能使用中微子束，也不能使用反中微子束，而需要使用两者的混合粒子束，否则就无法搞清离开 μ 子的符号

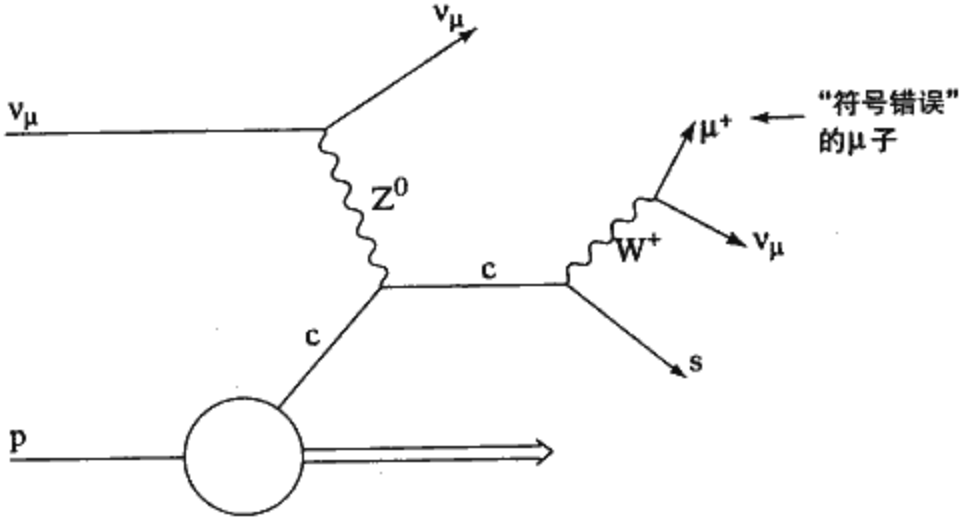


图 8.5　中微子散射能够通过“错误的 μ 子符号”表明质子中存在粲夸克，最终出现的是 μ^+ 粒子，而不是 μ^- 粒子

是否“错误”。

一般来讲，中微子实验是一项艰难的工作。人们无法沿着电子加速的轨道，产生和加速一束约束得很好的纯中微子束。中微子很难发生相互作用，这使得观测中微子束行为的实验很难进行。现代中微子实验使用亮度很大的粒子束，例如从高能质子束撞击固定靶的碎片中产生的粒子束，原始质子极高的速度转变为准直性相对较好的中微子束。另一种方法是合理利用巨大的靶-探测器一体装置，可以将观测到中微子行为的机会增加到最大。最近，人们已经搜集了上百万可用的数据，这足以满足很高的统计准确度要求。

正是精确的统计使得 CCFR 实验名列前茅。在最近进行的一项实验中，CCFR 合作小组将 800 吉电子伏的 Tevatron 质子打到铍靶上，产生含有 K 介子和 π 子的次级粒子束。这些飞行中的粒子在衰变过程会产生中微子。最终的粒子束是一种含有中微子的混合物，其中大多数都是 μ 子中微子，但反 μ 子中微子的量较少，其中还含有大量的电子中微子和反电子中微子，它们的能量都高达 600 吉电子伏。36 米长的探测器由 690 吨铁板组成，上面点缀着闪烁器和漂移室，仪器的后半部分是一台 μ 子谱仪，它可以用来探测和追踪 μ 子。一个 μ 子中微子撞击铁中的核子生成一个 μ 子，它能够被 μ 子探测器探测到。在过程中，从核子中撞出的夸克会变成强子，它们能够被铁板之间的活性层探测到。升级后的实验叫做 NuTeV（“Neutrinos at the Tevatron”——Tevatron 中微子的缩写），于 1996 年 5 月启动，实验使用了升级后的探测器以及一束中微子或者反中微子。实验者们通过电荷挑选出适当的次级 K 介子和 π 子，完成了实验。NuTeV 实验明确地显示出核子中存在粲夸克的踪迹，此外还发现了奇异海夸克的存在；核子中存在奇异夸克比存在粲夸克具有更加重要的意义。

夸克海：不对称且充满了胶

质子和中子可以通过交换上夸克和下夸克变成一样的粒子，换言之，它们通过同位旋对称性联系在一起。人们很自然地假设，质子中包含的反上夸克组分和中子中包含的反下夸克组分是相同的，即 $\bar{u}_{proton}=\bar{d}_{neutron}$，反过来也有 $\bar{d}_{proton}=\bar{u}_{neutron}$。但是，不存在一种对称性原理要求质子的反上夸克分布同它的反下夸克的分布相同；也就是说，没有任何证据说明 $\bar{u}_{proton}=\bar{d}_{proton}$ 是成立的。但是由于缺乏相反的实验证据，粒子物理学家们一直这样假设。

现在，物理学家们知道，这种假设并不正确。在 1998 年，费米实验室的 E866 合作小组的工作得到了令他们吃惊的结果，反下夸克的数量大大超过了反上夸克。E866 小组将 800 吉电子伏的质子打向静止的液体氘靶和液体氢靶，同时监测入射质子中的夸克同靶核子中的反夸克的碰撞，根据 Drell-Yan 机制，它们能够湮灭生成 μ 子对，见图 8.6。这两种不同靶的散射截面比率，可以直接转化为核子中反下夸克同反上夸克数量的比率，把该比率与已有结果比较，就能得到两者之间的差别。结果显示，在动

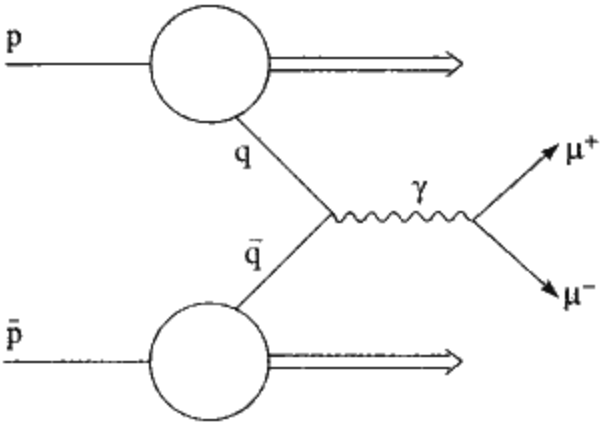


图 8.6 质子-反质子碰撞过程中的 Drell-Yan 机制。另外，反夸克也有可能从质子海中产生

量百分比较大的部分，两者贡献相等，但都极其微小。但对于动量百分比小于 1/50 的夸克和反夸克来说，反下夸克的数量明显超过了反上夸克的数量。

这是人们第一次通过核子海中的反夸克分布见证味道的不对称性。它也体现了另一种求和规则的破坏，该求和规则是由康奈尔大学的库尔特·古特佛里德（Kurt Gottfried）在 1967 年推导出来的，他那时正在寻找能检验基本夸克模型的简单预言。古特佛里德求和规则是说，电子与中子散射同电子与质子散射概率之间的差别，应该是一个简单的分数。这个定律的基础是夸克携带独特的分数电荷，并且质子的夸克组分是 uud，而中子的组分是 ddu。该定律还假定，核子内部要么不存在反夸克，要么至少核子中出现的反上夸克和反下夸克数量是相等的。CERN 的 NMC 合作小组在一个更严格的条件下，于 1991 年发现这个简单定律不成立，这些都为夸克-反夸克海存在的证据增加了可信度，并且还出现了一些令人吃惊的性质，理论物理学家们仍在解释夸克海的这些性质。

这里暗藏着一个假设，那就是 E866 实验以及 NMC 实验中，通过散射 μ 子在氢靶和氘靶中得到的反夸克分布是同一回事。总体来讲，实际情况是人们认为在像 E866 实验的那些强子-强子碰撞中，通过 μ 子、电子或中微子撞击核子得到的夸克、反夸克以及胶子的分布结果都可以互换。但是这掩盖了一项重要的假设。那就是强子之间的相互作用可以被完美地划分为几个没有关系的部分，第一部分是部分子的分布，第二部分是自由部分子的相互作用，最后是真实实验中观测到的碰撞产物，即粒子喷注。只有这样，利用一个实验中测量到的夸克和胶子描述其他实验才有意义。但是，这个被称做因子化（factorization）的假设并不总是成立。例如，在质子-质子碰撞中，质子束中的大量夸克很可能会使靶质子中夸克的分布改变，或者可能影响被撞击夸克离开碰撞

点的排列顺序。理论物理学家们同意确实存在这样的点，在该点处以上所述的现象会变得很显著，因此因子化假设就会失效。

但是，可以用这些分立的因子合理地描述像质子-质子碰撞这类相互作用，夸克和胶子分布实际上是普适的。从一个实验中得到的精确测量值可以在其他实验中被用做预测。这很重要。例如，理解 Tevatron 上顶夸克的产生以及估计在 LHC 有多少顶夸克诞生，都需要知道夸克和胶子分布。

就像在第四章提到过的那样，理论物理学家们对他们的顶夸克预言感到特别满意，其中一个原因就是在费米实验室中，顶夸克的产生主要来自夸克和反夸克湮灭。这说明，质子中含有的胶子相对不太重要，至少在费米实验室产生顶夸克的能量条件下是这样。这是一件好事，因为要想在核子中精确测量胶子分布，是一件比登天还难的事情，这对于预言底夸克在当前能量下的产生以及即将在 LHC 上进行的顶夸克诞生这样的过程都是很糟糕的消息，这两个实验都更依赖于胶子组分。问题的根源在于胶子不带电荷并且没有味道标签，因此通常用来发现夸克和反夸克的规则都不再适用。

但揭示胶子分布还是有可能的。例如，HERA 的实验者们就通过电子-质子碰撞所产生的喷注，测量了质子中的胶子海。在这些碰撞当中，从入射电子产生的光子会同质子中的胶子发生熔合，见图 8.7。每个出射夸克都碎裂成由 π 子和其他发生强相互作用的粒子组成的喷注。测量产生两束这种喷注的概率以及质子的残留物后，一旦算出了质子中夸克产生喷注的贡献，就能够得到质子中的胶子分布。HERA 所进行的第一次实验发表于 1995 年，那些实验是最早对占父质子动量百分比很小的胶子进行的直接测量，在该区域内胶子起支配作用。实际上，动量大于质子动量 1/10 的胶子寥寥无几，但是动量百分比较低的胶子数量非常多。这说明低动量百分比的部分子实际上都是胶子，这同用间接

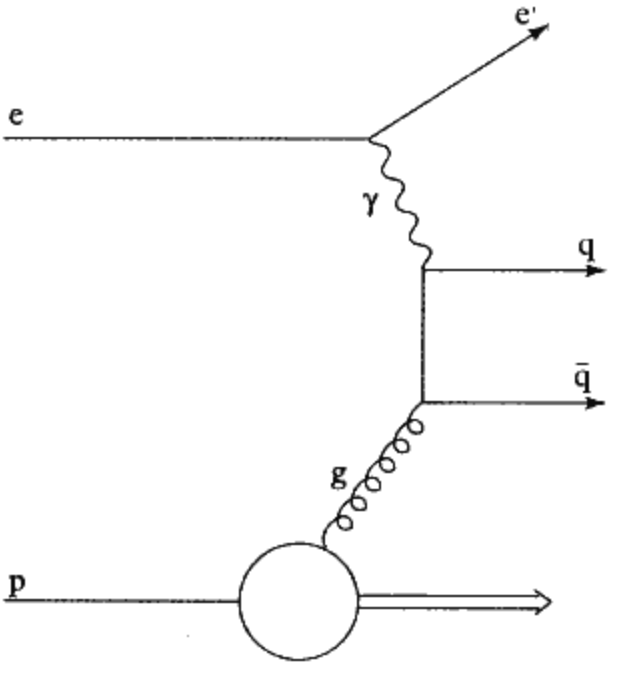


图 8.7 在电子-质子深度非弹性散射过程中的玻色子-胶子熔合，该过程能够探测质子的胶子组成

方法估计质子中胶子分布的结果一致。这个结果还具有直觉上的意义：如果胶子携带了大约 50%的核子动量，并且如果像其他实验显示的那样，占质子动量百分比很大的胶子数量非常少，那么在质子中必定存在大量的软胶子，见图 8.8。

柔软的海

按照 QCD 领域的先驱者、苏联理论物理学家尤金·列文（Eugene Levin）的说法，如何理解在低质子动量百分比区域，质子中胶子、海夸克同反海夸克的行为，是“QCD 理论中最艰难也是最重要的一个问题”。他曾经说，“至少在 QCD 理论成为一套在高能情况下描绘强相互作用的真正理论之前，这是一个应该被解决的问题”。大自然好像要再次强调列文的评论，HERA 所做的实验令人吃惊地发现，质子内部聚居着大量软夸克。

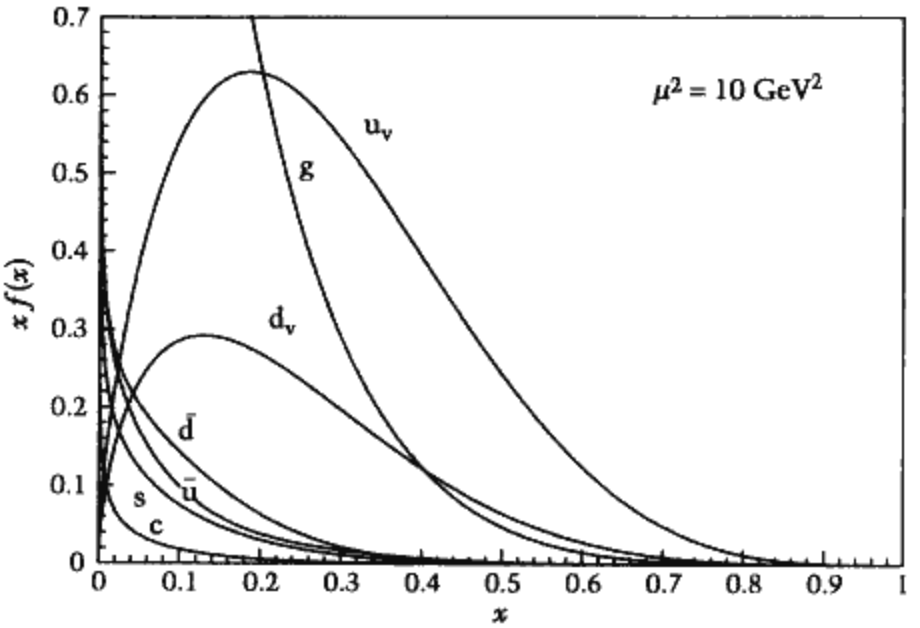


图 8.8 纵轴为实验测量到的质子中夸克和反夸克分布乘以动量百分比 x，横轴为 x。下角标“v”表示价夸克，“f”表示胶子或者特定夸克或反夸克味道分布。这幅图表示质子中存在少量但是可测量的奇异夸克和粲夸克。它还显示，质子中包含的反上夸克数目与反下夸克数目不相同。综合这张图中的夸克和反夸克分布，能够得到结构函数 F_2。图片来源：粒子数据小组

CERN 的实验者们利用 μ 子同固定氢靶和氘靶散射，探测到动量百分比为 0.01 的区域，观测到了一些新现象的蛛丝马迹。1992 年，也就是 NMC 小组发表研究结果的同一年，HERA 的实验者们在意大利的圣明尼亚托会议上发布了深度非弹性散射的初步数据。这些成果将夸克动量百分比值推进到了 0.0001，它只占在此之前最优秀的 NMC 数据的 1/100。质子研究领域从此揭开了新的篇章。

要撞出具有如此低动量百分比的夸克，实验者们需要用波长足够长的光子来探测这些波长较长的夸克，同时还要能够向被撞击的夸克传递许多能量。一种方法是制造出一束能量巨大的电子

束，用它以普通的方式撞击静止的氢靶。但是粒子束所需的能量太大了，我们无法实现；此外，由于散射电子的角度特别小，我们几乎无法看到它们，如果能观测到，那也是一项很难的任务。更好的方法就是建造一台对撞机，它不仅能使产生的碰撞产物容易被碰撞点周围的探测器观测到，并且能够在更加合理的能量下完成这项任务。这就是 HERA 所采用的方案。在 HERA 中，27 吉电子伏的电子束同 920 吉电子伏的质子束对撞，其效果相当于 55 000 吉电子伏的电子束撞击静止靶。碰撞点周围的 Zeus 探测器和 H1 探测器除了测量散射轻子之外，还可以记录发生强相互作用的碰撞产物，这同只测量散射轻子的传统方法有重要的差别。

不断追究结构函数图中前 1/100 的部分，看上去可能是一种不必要的迂腐做法，但是有许多人却希望看到图像中的这一部分被填满。毕竟这是大多数海粒子聚集的场所。从实用主义的角度来说，很多过程的计算，比如质子-质子散射生成 Z 玻色子和 W 玻色子、底夸克和顶夸克以及假定的 Higgs 粒子的过程，都依赖于对低动量百分比夸克和胶子数量的了解。检验求和规则要包括所有动量成分的贡献，因此也要求了解这些信息。理论物理学家们已经提出了各种各样的估计，设想图像是什么样子。他们的努力都集中于研究胶子问题上，因为通过胶子反过来就可以得到海夸克和反海夸克的数量，但他们的工作受到了数据缺乏的制约。

HERA 得到的“低 x”结果给一种叫雷杰理论（Regge theory）的前 QCD 理论制造了困难，尽管该理论有一些缺陷，但它能够很好地描述某些重要的散射截面，例如质子-质子或质子-反质子碰撞的总散射截面。雷杰理论是根据意大利理论物理学家图里奥·雷杰（Tullio Regge）的名字命名的，它用交换介子来描述强子散射，如果夸克的数量大致为常数，该理论就会更加自然。但是早在 1975 年，4 位苏联理论物理学家的工作就预计，夸

克的数量会有大幅度增长，他们分别是圣彼德堡的扬克·巴利特斯基（Yanko Balitsky）和列夫·利帕托夫（Lev Lipatov）以及位于诺沃西比尔斯克的布德科核物理研究所的维克多·法丁（Victor Fadin）和爱德华德·库纳耶夫（Eduard Kuraev）。他们为了囊括一层比一层更软的胶子贡献，提出了一套描述方法，即所谓的 BFKL 方程，利用该方程可以得到胶子数目随动量百分比变化的情况。另一套方案是 DGLAP 演化方程方法（它的名字是由 Yuri Dokshitzer、Vladimir Gribov、Lev Lipatov、Guido Altarelli 和 Giorgio Parisi 的名字取第一个字母缩写得来的），该方法允许夸克和胶子之间发生相互转化，将其应用到胶子数量问题上，能够描绘出胶子数目怎样随着探针波长的变化而变化。

这两种方法都不可能凭空预言胶子的数量，它们在一开始都需要提供一些测量值。但是如果选取适当的起始分布或者基本模型，这两种方法都能够符合新得到的数据。但是仍然存在一种困难，因为所有预言的分布都是关于胶子的，而一般得到的却是夸克组分的值，两者之间还有一定的距离。直接测量胶子本身分布的技术不断改进，使我们可以更加精确地测量胶子分布，从而有助于检验理论物理学家们的模型，例如 HERA 通过粲夸克-反夸克对产生或 J/ψ 粒子产生的方法来测量胶子分布。同时，理论物理学家们正在努力理解两种方法之间的联系，它们在某个层次上必定能够很自然地相互融合到一起。理论物理学家们还必须面对 BFKL 模型的预言，即胶子数量会激增，这会破坏幺正性。为了克服这种困难，他们提出一种饱和思想，即随着质子中的胶子数密度越来越大，在密度值达到某个点时，胶子之间会发生相互作用并相互结合。这种胶子再结合的机制限制了质子内部所包含的胶子数量。最初的估计认为，只有在超出 HERA 观测范围的层次上，才能产生饱和现象。但是有些人认为，低动量百分比的胶子会挤在价夸克周围，形成一个“危险区域”，HERA 可能会观

测到它。

夸克与核子

大自然还为核子内部的运行机制附加了另一个难题：作为原子核一部分的核子中的夸克，其分布同自由核子内部的夸克不同。换句话说，核子内部的夸克对核子所在的环境很敏感，就好像它们能够通过某种方式感知到邻近核子中的夸克和胶子。这就是所谓的 EMC 效应，该效应是根据 CERN 的欧洲 μ 子合作小组（European Muon Collaboration）的名字命名的，该小组在利用铁靶和氘靶进行深度非弹性散射实验、测量核子的结构函数时，首次观测到这种差别。该小组最初的结果发表于 1983 年，随后该小组以及其他一些小组又对该实验进行了改进，并研究了更多的原子核。他们的结果表明，这种差别以比较复杂的方式依赖于被撞击夸克的动量百分比，见图 8.9。此外，对于从氦到铅之间的所有元素，结构函数的畸变模式同重核很相似，并且随着核子的整体密度不同而有所不同。

最初，理论物理学家们寄希望于找到某种能解释新结果的简单规律。但是，他们再一次受到幽灵一样的非微扰 QCD 理论的困扰。他们根据 QCD 理论创造了各种模型，试图解释在动量较高的一端夸克分布骤然下降的现象。一些人认为，原子核是许多核子簇的集合体，每个簇中都有许多三夸克结构。另一些人认为，对于较重的原子核禁闭的区域——也就是束缚胶子与夸克的牢笼的大小——会延伸到核子之外，甚至超出原子核自身的边界。根据传统的观点，介子会将核子束缚到一起形成原子核，这些介子也会占用一些夸克的动量，在核子中一般能够发现这些介子。要解释低动量下重核散射界面的下降现象还需要加入一些更多的内容，见图 8.9 中低 x 区域中曲线的下降。这种所谓的核屏

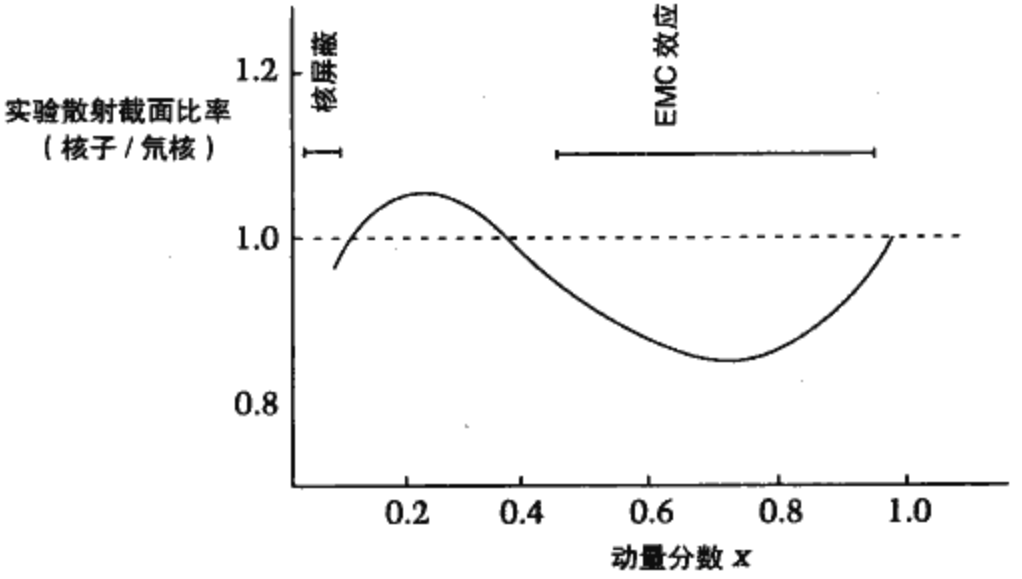


图 8.9 EMC 效应以及原子核屏蔽（shadowing）图。纵轴是原子核散射截面同氘核散射截面的比率。如果原子核中的核子对探测光子的反应同氘核中的核子相同，图将如水平虚线所示

蔽（nulear shadowing）是由于邻近核子中发射出低动量、长波长的胶子，这些胶子之间也发生相互作用。这是胶子再结合的一种方式，这种效应在单独的核子中会引起胶子饱和，前一部分已经提到过这一点。

屏蔽效应和 EMC 效应提出的问题，就是要找到一种能同时描述原子核结构函数和核子结构函数的全面 QCD 表述方式。这样就能够把对核子结构的认识转变为对原子核结构的认识。至少从原则上讲，QCD 理论是一种能够解释所有原子核结构的基础理论。但实际上，理论物理学家们却要利用各种剔除了 QCD 理论中错综复杂的内容之后的核结构模型来理解原子核。哈罗德·弗里茨写道："讽刺的是，QCD 理论没有提供多少有关强相互作用

的内容，而这正是我们研究的最初动机。”根据维尔茨克的想法，QCD 理论与核物理之间的关系同 QED 理论与化学的关系很相似。完善的电磁相互作用量子场论构成了化学反应的基础，但是在现实中，许多化学家都用键、轨道等一些概念进行工作。

实际上，化学为 QCD 理论怎样用于核物理提供了一套生动的例子。氖是用于霓虹灯的一种材料，它在室温下是一种气体。它真是一种模范元素，它的原子有适当的电子数，使它呈很强的惰性。但是在－246℃，氦会变成一种液体，温度再低 2℃，它就会变成晶体。这就表示，尽管每个原子都是一个不发生相互作用、不带电荷的近乎完美的球体，但各个原子之间实际上还是存在某种剩余作用力。要想理解这种力的起源，就要牢记每个原子实际上并不是无内部结构的坚硬小球，而是由原子核和电子组成的。实际上，当两个原子相互靠近的时候，它们能够并且确实感受到彼此所带的电荷。每个原子所带的电子云相对于原子核都发生了微小的偏移——突然间，每个原子带正电的核心不再同电子云的中心重合。结果，每个原子都会呈现出一点儿电极性，因此距离很近的两个原子之间会存在吸引力。这就是范德瓦耳斯力，这种力根据荷兰物理学家约翰尼斯·范·德·瓦耳斯（Johannes van der Waals）的名字命名。

如果把这种模型应用到核物理中，电荷为零的氖原子就要替换成颜色荷为零的核子。核子之间不会直接有基于颜色的作用力，就像氖原子之间不存在直接基于电荷的作用力一样。但是相邻的核子能够改变彼此颜色云的分布，其方式类似于氖原子之间的相互影响。同范德瓦耳斯力引起了氖原子之间的相互作用力类似，变形的颜色云之间也会产生一种类似的力。正是这种类似于范德瓦耳斯力的相互作用，在原子核中将核子结合在一起。

发生强相互作用的夸克中心发生偏移，导致产生将单个核子结合到一起的核力。QCD 理论最终可能会为我们提供理解原子核

的基础，但借用维尔茨克的话说，QCD 理论在原子核结构实际问题中的角色，基本上同“圣水”相同。

经验教训

通过深度非弹性散射研究质子的结构，不仅帮助粒子物理学家们理解了质子怎样结合在一起，还为他们提供了一种检验 QCD 理论的实验台。其中最好的一种检测方式，就是观测核子中夸克分布的模式，这种模式实际上同最简单的模型有所差异。

在深度非弹性散射过程中，探针光子会观测到点状夸克，用行话来讲，这实际上就是简单的部分子模型，这就是故事的结局。但是随着探针光子波长的降低，它能够看到一个夸克和夸克发出的一个胶子，或者在更细微的层次上能看到一个夸克和一个夸克-反夸克对，等等。探针波长越短，探针光子就能够看到更多粒子。随着粒子数的增加，父核子动量也必须散布得更广，找到一个具有特定质子动量百分比的夸克的可能性也会发生改变。经 QCD 理论改良过的部分子模型说明，存在大量低动量的组分，并且结构函数确实以一种特有的方式依赖于探针波长。

理论物理学家们利用马上就要被撞击的夸克发出的胶子来分析该现象。夸克的动量最初是父质子所带动量的一部分，由于胶子带走了一部分动量，它的动量会变得更小。因此由于存在夸克发射胶子的可能性，探针光子遇到具有特定质子动量百分比夸克的可能性也发生了变化。关键点在于，探针光子的波长越短，夸克辐射胶子的可能性就会增加。这说明，根据 QCD 理论，具有特定质子动量百分比的夸克分布测量值，实际上依赖于探针光子的波长。人们最初得到的经验是尺度不变性，即结构函数依赖于某些无量纲的量，这反映了散射体本身是点状的事实。这是一种需要修正的经验：尺度不变性遭到了破坏，但是是以 QCD 理论

描述的、特别的方式。尺度不变性支持了点状的散射中心观点。

发射胶子仅仅是在质子的夸克分布中，尺度不变性被破坏的一个原因。产生的胶子也会发生作用，它们还会产生更多的胶子，或者胶子会产生更多的夸克-反夸克对，它们反过来都会成为被撞击的部分子。结构函数会随着探针波长的变化而渐变，也就是慢慢变化。这种变化可以在考虑各种夸克发射胶子、胶子发射夸克以及胶子发射胶子的基础上，利用微扰 QCD 理论进行计算。最终得到的 DGLAP 演化方程，已经在质子中软胶子数量的讨论中出现过了。

人们或许会吃惊地发现，尺度不变性破坏的想法诞生得比 QCD 理论本身还要早。例如，在 1973 年，格罗斯和维尔茨克发表的那篇提出规范理论中渐近自由的伟大论文中，就说明了我们现在所知的、由 QCD 理论预言的尺度不变性破坏现象。不久之后，在费米实验室一次 μ 子同铁靶的深度非弹性散射实验中，尺度不变性破坏的实验证据也出现了。5 年之后，CERN 的实验者们利用中微子的手性，测量了宇称破缺的第三结构函数中尺度不变性的破坏，从而在具有细小差别的环境中为 QCD 理论提供了支持。尺度不变性破坏在夸克碎片转变为强相互作用粒子的过程中起着重要作用，测量尺度不变性破坏现象是一项很大的工程，实验者们从各种不同的实验中收集了大量数据。尺度不变性的破坏不仅仅能为 QCD 理论提供检验。随着探针波长的降低，结构函数会向动量百分比低的区域逐渐倾斜，这同核子内部的胶子分布有直接关系。实际上，反过来考虑，尺度不变性破坏现象的测量会提供一种得到胶子分布的方法。

理论物理学家们不能用微扰 QCD 理论凭空计算出结构函数，但是他们却能计算出结构函数的演化方式；结果并不像完全尺度不变性预言的那样，结构函数与探针波长无关。尺度不变性破坏的程度要依赖于耦合强度，因此也与 QCD 理论的特征数字——

尺度参数 Λ 有关。这些都使得在深度非弹性散射过程中看到的尺度不变性破坏，成为已有的检验 QCD 理论的最好方案之一，甚至有些人认为它就是最好的。

下面是比较理论和实验的一种标准方法：先测量出特定探针波长下的夸克分布。DGLAP 演化方程改进后可以计算更多的复杂费曼图，利用最开始测量到的分布和该方程可以预言不同探针波长下的夸克分布，然后拿这些结果同实验比较。要么能够从其他地方引入一个尺度参数值，要么可以设计出一套拟合程序，从而拟合实验数据得到一套最合适的参数，其中也包括尺度参数。

在许多不可避免的复杂性当中，最微妙、影响最广、最令人感兴趣的，是由额外胶子交换引起的“基本”过程末态变化。例如，图 8.10 中的图（a）所示的深度非弹性散射过程，一个额外胶子将质子碎片同被散射的夸克联系起来。在图 8.10 的图（b）中，显示了一次质子-反质子散射的 Drell-Yan 过程，质子碎片同反质子内部的一个反夸克联系在一起。由于要涉及一种神秘的数学方法，这一类型的过程被称作“高扭度”（higher-twist）效应。

这是一个时间问题。例如，当一个波长很短的光子刺入质子，一切都发生得特别快。因为波长短的虚光子寿命也很短，这是由不确定性原理所决定的。根据所谓的脉冲近似理论，一个被短波长光子撞击的夸克离开得很干脆。但是当相互作用并不是发生得特别快、光子波长和时间尺度都比较长，就有时间发生额外的相互作用。这说明正常情况下只作为观测粒子的部分，会和散射过程以某种方式联系起来，也就是说发生了关联。抢劫银行的劫匪逃跑，跑的时候就会同旁观者有（高扭度）相互作用。

高扭度效应不能通过微扰理论计算，因为从本质上讲，这种效应在长波低能端更为重要，而这正是微扰论的可信度遭到质疑的范围。正因如此，并且有特别多的过程涉及高扭度效应，才使得它们如此富有挑战性且引人入胜。理论物理学家们只能将某些

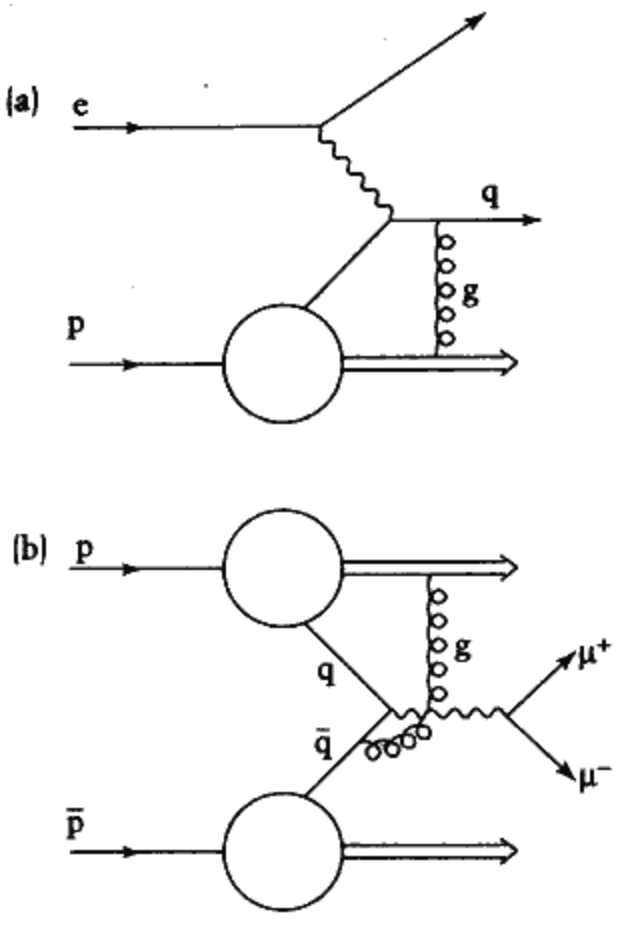


图 8.10 高扭度问题包含了同观测粒子之间的额外相互作用，图 (a) 是深度非弹性散射过程，图 (b) 是质子-反质子对撞过程

特定过程的高扭度贡献估计到某个量级，或者根据结构函数数据对它们进行模拟，或者至少能够确定在哪些实验中高扭度效应的影响最小。

先撇开高扭度效应这类复杂的情况不管，尺度不变性的破坏仍是检验 QCD 理论的一个重要的实验台。弗里茨在 1992 年写道，“观测到的尺度不变性破坏现象同 QCD 理论预言得如此一致，使我们不得不由衷赞叹理论的优美与成功”。

用光击打质子

将“深度非弹性散射”中的“深度”拿走，就会发生一些令人吃惊的事情。当你用一个电子轰击质子，刺探它的内部，你可以调整实验参数，使传递碰撞的光子变为真实光子，或者至少非常接近真实光子，即所谓的“准真实”光子，它们不再是虚光子。例如，当被散射的电子离开碰撞点时不偏移，就会产生真实或准真实光子。真实或准真实光子参与的散射过程称作光致产生(photoproduction)①。

就这样，潘多拉的盒子被打开了。在光致产生过程中，非微扰效应普遍存在，用高扭度方法理解问题是不可避免的，比约肯尺度不变性变得毫无意义，质子的夸克组分也几乎无法看到。但是，光致产生过程必须同深度非弹性散射现象天衣无缝地衔接起来。这一切都使得光致产生过程非常吸引人。

光子也会发生一些奇特的事情，可以通过一种奇怪的方式用它来理解质子的结构。一个光子周围会伴随着一团虚粒子，像电子、正电子等，甚至还有夸克和反夸克，它们都按照量子世界的规则从真空中诞生，又回到真空中去。对于一个真实光子或者近似真实的光子，这种量子旅行的持续时间大于 10^{-25} 秒，伴随光子的虚夸克和虚反夸克也是这样，因此它们产生的胶子有时间参与强相互作用。简言之，光子是一种典型的基本粒子，但它的行为并不像一种简单的基本粒子。它看上去包含着夸克和反夸克，因此看起来像是同质子一样的复合粒子。因此光子也能够参与强相互作用。在量子世界中，任何事物都同以前看上去的样子不同——我们要更加仔细地观察，它们的面貌是时常变化的。

---

① 译者注：或简称光生。

在某些过程中，光子的行为是一个不可分割的整体，这种过程称作直接光子相互作用，见图 8.11 中的图（a）。与之相反，通过周围的夸克与反夸克发生相互作用的光子称作分解光子，见图

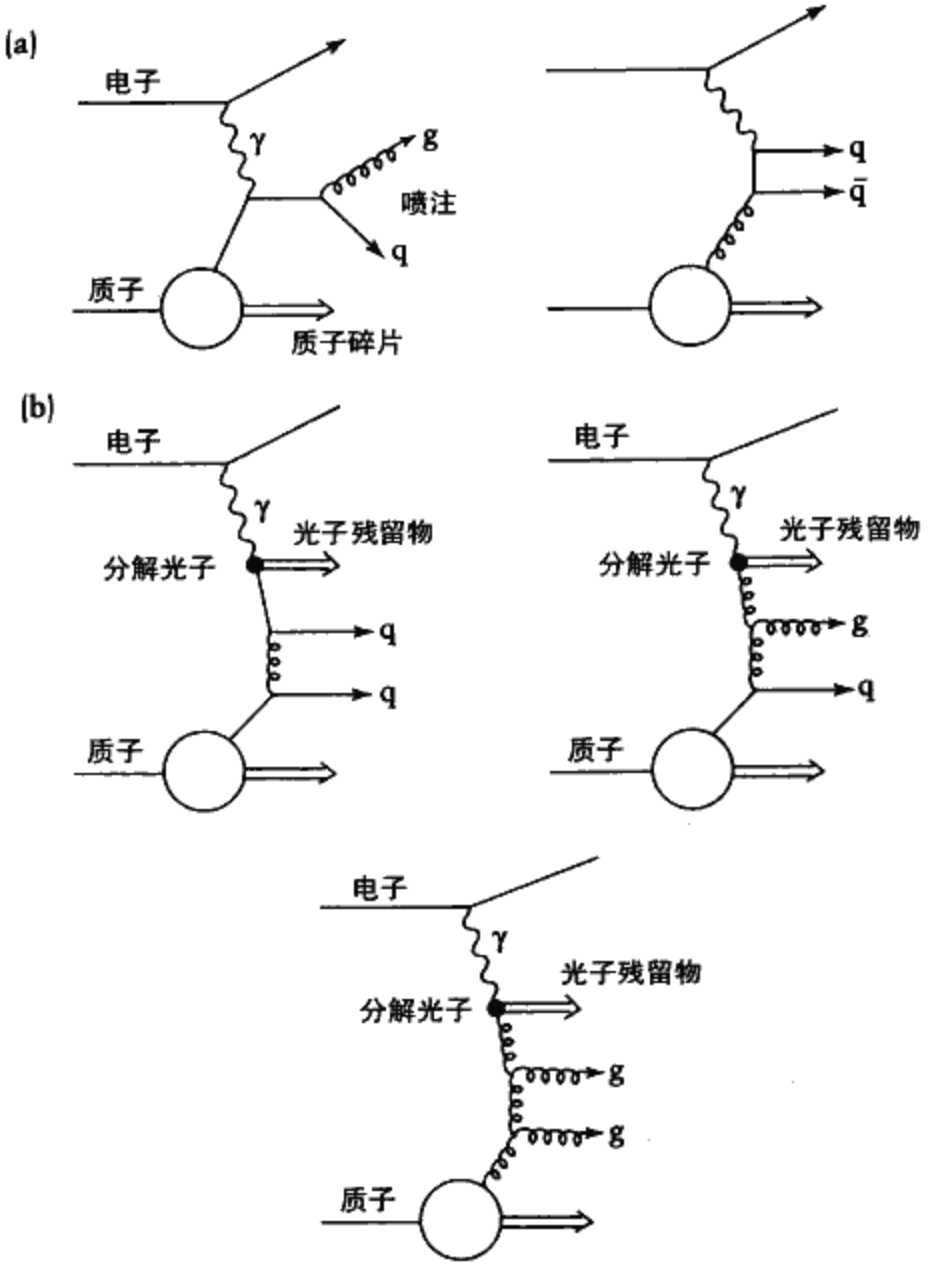


图 8.11 光子-光子散射过程中的直接光子［图（a）］同分解光子［图（b）］。图（b）中涂黑的小点表示分解光子，它能够产生光子残留物

8.11 中的图（b）。在某种意义上讲，分解光子就是一种光子碎片。

如果一个分解光子通过夸克或反夸克发生相互作用，就说明光子“内部”存在某种夸克和胶子的分布。换句话说，同核子的结构函数类似，应该存在一个光子结构函数。但是应该怎样测量这种东西呢？答案就是按照核子的例子进行测量：取一个光子，然后把它放到显微镜下进行观测。在实践中，这就表示进行深度非弹性散射，利用电子束产生的虚光子探针，检验真实或者准真实光子。这听起来或许不可思议，但执行起来却十分简单。只需要将一束电子同一束正电子碰撞，对实验进行调整，就可以观测到（比如）电子束中的虚光子，同另一束粒子——正电子中的真实或准真实的光子——以普通深度非弹性散射的方式进行碰撞，见图 8.12。这时，虚光子探针就能够看到目标光子分裂为夸克和反夸克。

这类过程在 1981 年首次被 PLUTO 小组用 DORIS 电子-正电子对撞机观测到，稍后又在 PEP 上观测到，在 LEP 和其他电子-正电子对撞机上也曾经对它们进行过研究。光子确实有结构函数；换句话说，光子中所包含的夸克成分真实并且可以测量。LEP 得到的结果证实了夸克的分布图会稍微鼓起一点，大多数夸克的动量占父光子动量的百分比都比较居中。这正是人们预料之中的，因为整个过程都必须不断从单个光子产生夸克和反夸克，光子的动量在夸克中间进行简单的分配。尽管随着探针波长的降低，光子中可见的夸克组分增加得很慢——这符合 QCD 理论预言，但人们还不清楚低动量百分比区域的情况。理论物理学家们推测，夸克分布在低动量百分比区域会呈现剧烈的上升趋势，就像在质子中观测到的一样。但到目前为止，在动量百分比达到 1/400的层次上，还没有任何出现上升的迹象。

研究光子结构函数的一个诱人之处在于，光子看上去很简

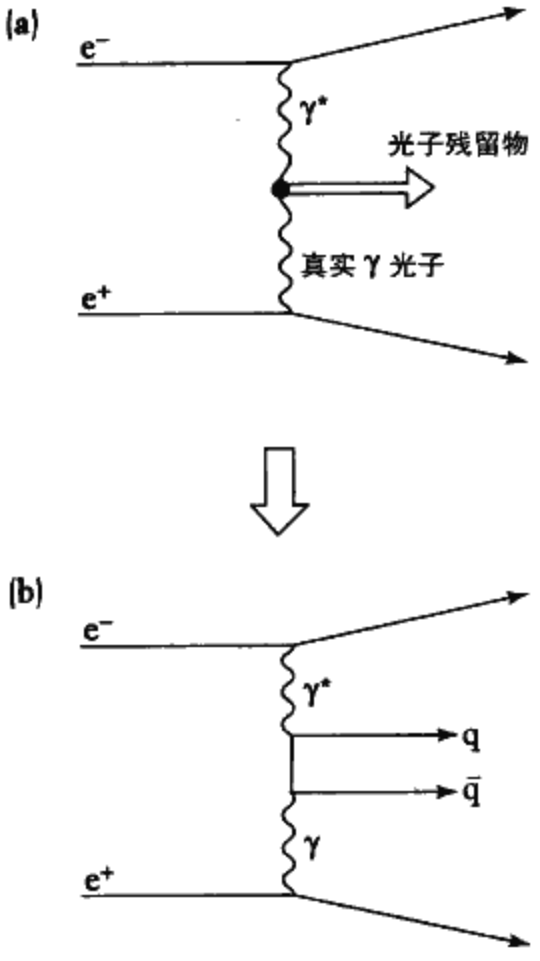


图 8.12 用深度非弹性散射测量光子的结构函数。一个入射电子产生一个虚光子探针 γ^*，它会同来自于正电子的分解光子相互作用[图 (a)]。图 (b) 是分解光子的简单贡献

单。在 1977 年，普林斯顿大学的理论物理学家爱德华·威腾（Ed Witten）证明，利用 QCD 理论可能（至少在理论上）预言光子内部夸克分布的实际情况。这除了能显示出光子的夸克组分怎样随着探针波长变化而变化之外，还包含一些内容：不存在相应的方法仅根据“纯想象”就能推导出核子所含的夸克情况。差别在于，对于光子而言，整个故事有一个很简单的开端，用费曼图的语言来讲就是光子-夸克-反夸克交叉。但质子却没有如此简

单的起始点。

粒子物理学家们曾经希望，光子结构函数能为他们提供一种干净利落的方法，来测量 QCD 理论中的耦合强度，这种方法在威腾关于光子中夸克分布的计算中以一种简单而美妙的方式体现出来。但人们的希望落空了，这种希望在实验中毫无立足之地。实际情况表明，通过纯理论预言光子中夸克分布的测量值行不通。但是，人们还是对光子比较感兴趣，因为它是一种比核子简单的物体，理解了前者的强子结构就会为研究后者开辟道路，这就是原因。

当真实或者准真实光子同质子发生相互作用时，电子散射实验告诉我们，光子中的夸克会让光子患上精神分裂。光子们能够直接相互作用，也可以作为夸克和反夸克源通过夸克来相互作用，这两种情况都必须考虑。同复杂程度相似的直接光子相互作用作比较，分解光子的相互作用会产生一束额外的喷注，它的方向接近入射光子的方向，见图 8.13。对于 HERA 中的光致产生过程，质子残余喷注会沿着质子初始方向出射，而分解光子相互作用产生的光子残余物却沿着散射电子的路径行进。最简单的直接光子过程和分解光子过程，都会导致两束相对于碰撞粒子束的大偏角喷注产生，即所谓的“高横向动量喷注”。直接光子过程和分解光子过程之间的差别在于，后者还会产生额外的光子残余喷注。这两种类型的事件都能够在 HERA 中观测到，在适当的条件下，分解光子散射发生的概率比直接光子过程大。在 HERA 中，光子残留物可以清楚地观测到，但在较早的固定靶实验中却不行，因为它们同粒子束极为相似。

真实光子发生的深度非弹性散射以及真实光子同质子发生的碰撞过程，都只涉及一个真实光子。但一个真实光子同另一个真实光子碰撞，并且产生新物质的过程也有可能发生。QED 理论预言了这种过程的存在，但人们直到 1997 年，才第一次在实验中

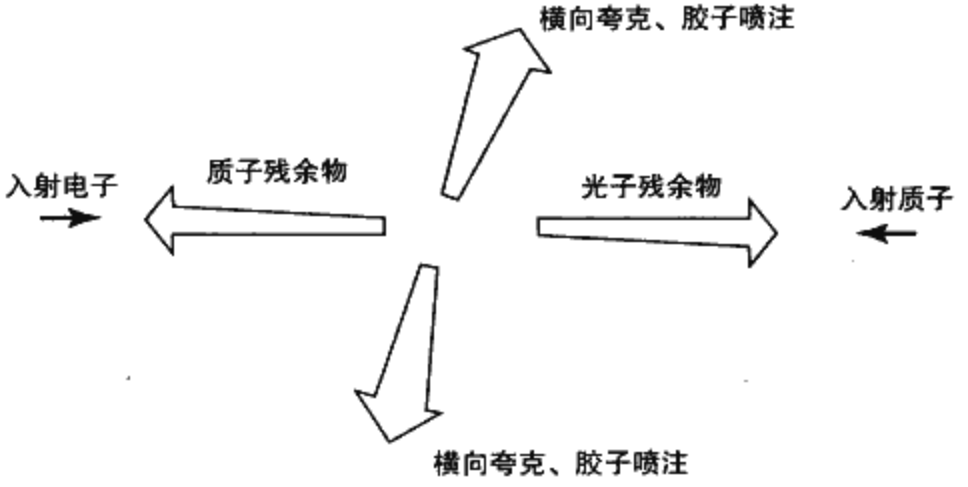


图 8.13 质子同分解光子碰撞，得到四喷注事件，质子和光子都产生了残余物，还有由夸克和胶子形成的横向喷注。在 HERA 中，由于质子的动量很大，喷注会偏向于质子的方向

通过光子-光子散射产生物质，在 SLAC 上工作的一个小组无可辩驳地用真实光子制造出了电子-正电子对。该小组通过用一束非常强大的激光迎面轰击用 SLAC 加速的电子束得到了该结果。迎面而来的电子将光子沿原路打回去，该过程增加了光子的能量并将光子变成了 γ 光子。这些 γ 光子会同入射的激光束光子碰撞，从而产生电子-正电子对。

没有人怀疑该过程，但如果确认一下真空真的可以这样摩擦出“火花”会更加保险一些，该小组的一个成员——艾德里安（Adrian Melissinos）做了这项工作。5 年之前，TRISTAN（日本电子-正电子对撞机）的 AMY 合作小组独自完成了一次小型真空点火试验。该小组所用的光子是差不多真实的那种，因此与 SLAC 的真实光子实验稍微有点差别，但该小组能够用光子-光子散射从真空中拉出强子。AMY 合作小组让电子和正电子对撞，两种粒子的散射角都很小，这样传递相互作用的虚光子就会近似

变成实光子，见图 8.14。AMY 合作小组发现大量强相互作用粒子从碰撞点发射出来，只有发生碰撞的一个甚至两个光子产生夸克或反夸克，才能解释该现象：他们的实验是第一次在非电子-光子深度非弹性散射实验中确凿地观测到分解光子，也是光子含有胶子的第一个信号。AMY 合作小组以及 TOPAZ 小组和各个 LEP 合作小组，那时已经能够区分由准真实光子碰撞产生的强子喷注：这种过程总共会出现多达 4 个强子喷注，每束喷注分别是由两个“基本”夸克、胶子或反夸克产生的，其中有一两束来自于光子碎片。

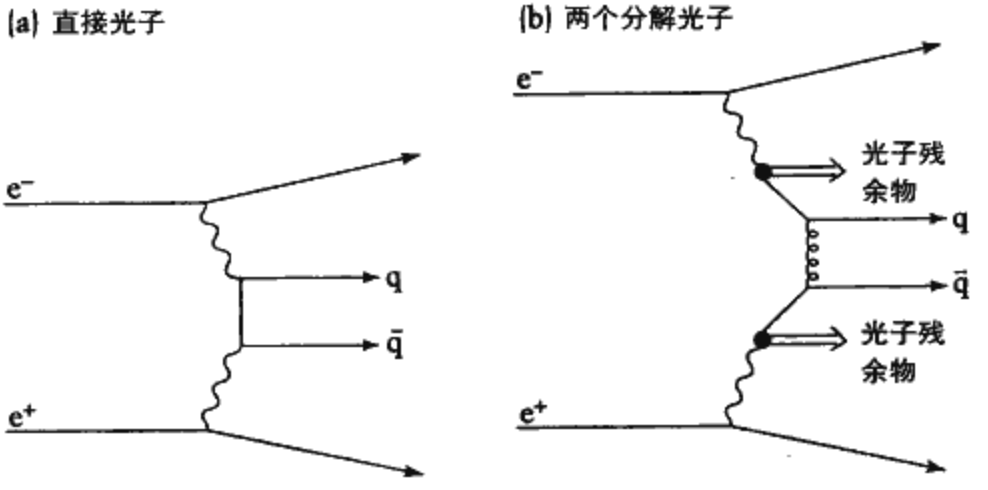


图 8.14 “光子-光子”散射过程，能够通过两个直接光子［图(a)］、两个分解光子［图(b)］或者一个直接光子和一个分解光子（没有表示出来）实现。这三种情况分别能够产生二喷注、四喷注和三喷注事件

将两个准真实光子撞到一起，甚至有可能摩擦出质子-反质子对。这并不是异想天开——这种现象已经在几个实验中观测到了，例如 CESR 的电子正电子对撞实验。光致产生就这样带来了量子世界的另一个奇迹，可以从“简单”光子的碰撞碎片中构造像质子这样复杂的粒子。从真空中可以毫不费力地变出一个完整

而美妙的氢原子核，这是多么复杂而壮观的景象，我们就生活在这样一个宇宙中。

自旋：一次危机

大自然偶尔会提供一些免费的礼物。利用 HERA 进行研究的天才实验者们，试图用这种礼物解决粒子物理中的一个巨大谜团——核子自旋的起源。

在存储环中运动的电子或正电子会发出同步辐射。作为一种副产品，环中的电子会变得横向极化。换句话说，它们的自旋方向会垂直于它们的飞行方向。两位苏联物理学家索科洛夫（Arsenii Sokolov）和艾格·特诺夫（Igor Ternov）早在 1963 年就首先发现了这种现象，他们两人都来自于莫斯科州立大学。自旋能整齐地排列，是因为那些自旋反转后与磁场方向反平行的电子更容易发射光子。这种自旋的同向排列是大自然的免费赠品，但是直到最近，加速器物理学家们才开始利用它。可是问题在于，如果要真正利用它们，电子或正电子的自旋应该指向飞行的路径，即所谓的纵向极化，而不是上面提到的横向极化。

在 HERA 上做过一个叫做 HERMES（HERA 自旋测量实验，或者对于狂热的希腊神话迷来讲，它是指宙斯的儿子、神的使者“赫尔墨斯”）的实验，实现了这一要求。HERMES 是为了研究自旋相互作用而设立的。与 H1 探测器和 ZEUS 探测器都建造在质子-正电子碰撞点上不同，HERMES 中只有正电子通过，观测的是正电子和 HERMES 中定制的固定靶之间的碰撞。探测器上游的特殊磁铁组合改变了入射正电子的自旋取向，因此它们到达靶的时候全部变成纵向极化。另一组磁铁组合，即自旋旋转器，在靶的下游将自旋方向再改回去。这样，这些正电子就会绕着环继续它们的旅行，不发生任何改变。这是人们第一次在如此

高能的环状加速器上制造纵向极化电子。完成这项任务可不是一件轻而易举的事情：近乎完美的自旋一致性要求在整个 HERA 环 6.4 千米长的圆周上都要安装聚焦磁铁，一些人甚至怀疑过这种技术的可行性。第一组数据在 1997 年向世人公布，从而证明了它确实可以实现。

该数据还证实了 HERMES 开拓的另一种新奇技术的可行性。实验所用的靶是位于粒子束路径上的一团气体。气体不需要容器承载，否则粒子束和某些容器材料碰撞会扰乱真实信号。气体通过管道输送到靶区域，随后又被功率很强的泵抽走，这样气体就不会弥散于粒子束管道、破坏真空了。

最早被用作靶的气体原子是极化氦-3 原子（^{3}He）。让这种气体原子变为实验所需的极化原子核，是一项极其精细的工作。在 HERMES，实验者们用一种很弱的射频场，将氦-3 原子激发到一个亚稳高能态上，处在该态的原子中的电子，不再成对地呆在能量最低的结构中。通过用极化激光照射这些原子，使它们极化后，再让处于激发态的原子同另一些作为靶原子的氦原子碰撞。在碰撞过程中，激发态原子中电子的有序自旋运动会传递给作为靶原子的原子核。最后实验上用的就是这些靶原子，见图 8.15。

相对于那些用固体丁醇或固体氨等较传统的固体靶方法来讲，用氦-3、氘或者氢这样的气体靶，一个重要的优点是最终得到的信号非常清晰。例如，在氨中有 17 个核子，但其中只有 3 个能被极化，因此探测器中所测得的大多数事件，都是同靶的非极化部分碰撞造成的，因此会稀释有效信号。对于氢和氘，所有核子都是极化的。而对于氦-3，原子核中的 2 个质子配对，形成了一个自旋为零的结构，唯一的中子会发生极化。由于不存在简单的中子靶可用，尽管三个核子中只有一个能够极化，极化的氦-3 原子核仍然是最好的替代品。HERMES 和其他两个 SLAC 实

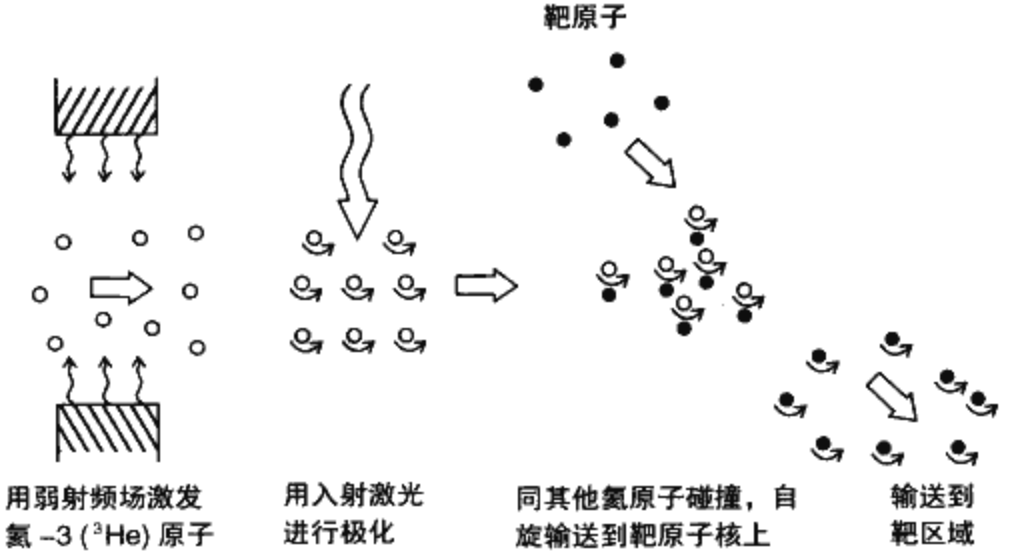


图 8.15 在 HERA 中产生极化的氦气体靶的过程。为了清楚起见，那些用作靶的原子都涂黑了

验——E142 和 E154 都利用了极化氦-3 气体靶。

要完成一个典型的极化实验，靶的自旋同探针粒子束的自旋要先排列成同一方向，然后指向相反的方向。对于 HERMES 实验，该要求通过反转制造靶原子的极化激光束，从而反转靶自旋的指向来实现。在两种情况（自旋平行或反平行）下，测量到的散射截面差反映了靶核子中自旋组分的情况，见图 8.16。

那么，HERMES 到底发现了什么？根据 HERMES 得到的结果，上价夸克的自旋方向同父质子相同，而下夸克的自旋相反。但还存在不足之处：HERMES 小组证实，只有 25%～30%的质子或中子自旋是由内部夸克引起的。根据理论，这个数字更有可能是 60%，甚至 100%。这种偏差不能令人满意。实际上，就像最开始看到的那样，这个结果令人大失所望，因此在粒子物理学中诞生了一个新词：自旋危机。

自旋危机的名字现在有点过时了，因为这种危机已经持续了太长时间，已经丧失了危机这个词所带来的紧迫感——这种病症

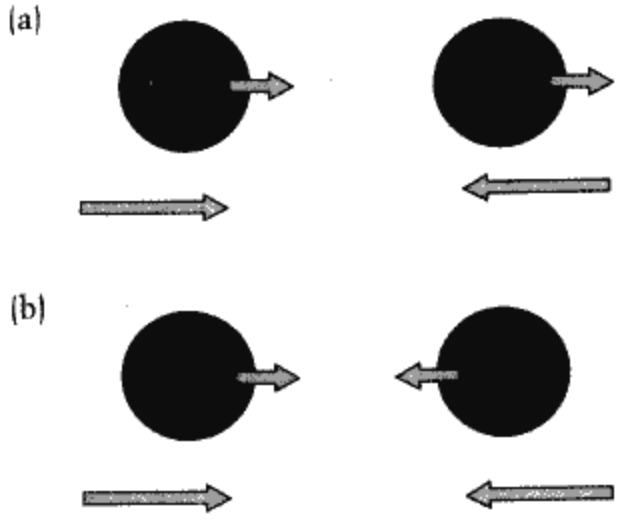


图 8.16 粒子与自旋平行的粒子［图（a）］碰撞以及与自旋反平行的粒子［图（b）］碰撞情况的对比，揭示了靶核子中自旋组分的情况

是慢性的，而不是急性。但其核心问题——怎样解释核子的自旋仍然存在着，虽然近来人们给它起了一个更好的名字，叫做“质子自旋问题”。但在 1988 年，这个问题迫在眉睫，自旋危机的说法非常贴切，那时 CERN 的 EMC（欧洲 μ 子合作小且）公布了一项用极化 μ 子同极化质子散射的实验。EMC 的发现——“总夸克自旋只占核子自旋很小一部分”——给人们带来了巨大震惊。此外，EMC 得到的结论同 1983 年进行的 SLAC 实验不一致，那次实验的范围有限，并没有显露出任何令人吃惊的征兆。

本来期望得到某个很大的数，比如 60%或 100%，但实验却给出一些很小的数值，而实际的统计结果是零，这并不是一个能让所有科学家甘心接受的结果。CERN 的 μ 子自旋合作小组跟随 EMC 的脚步继续前进，他们承担着艰巨的任务，即弄清自旋问题的真相。同 EMC 一样，他们利用了大自然提供的另一份免费午餐——π 子。利用 CERN 的 SPS 产生的质子撞击铍靶，能够产生 π 子，这些 π 子通过弱相互作用衰变为 μ 子，这些 μ 子天生就

是极化的，是一种自然的纵向极化探针，见图 8.17。同时，SLAC 的几个实验者组成的小组利用极化电子探针来探索自旋结构。SLAC 的数据更加精确；而 CERN 的数据却能够涵盖更大的范围。特别是 SMC 能够达到更小的 x 值，最低的 x 值可以达到 0.003，这个数字表示被撞击的夸克仅仅带有 0.3% 的父核子动量。

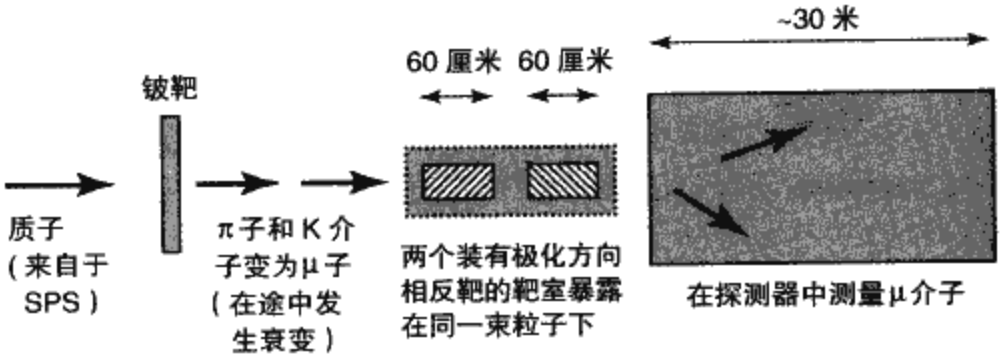


图 8.17 SMC 进行的极化 μ 子-极化质子实验示意图。质子撞击铍产生的 π 子和 K 介子衰变可以得到自然极化的 μ 子。两个极化方向相反的靶同时暴露在 π 子束下，这两块靶冷却到 50mK（0.05 开尔文）

SLAC、SMC 和 HERMES 实验都是深度非弹性散射实验：一个轻子被靶弹射到探测器上，探测器会测量出轻子的能量和运动轨迹。但是这次，靶粒子和探针粒子——电子、正电子或者 μ 子——都是极化的。实验的目的并不仅仅是总结出结构函数——核子中散射中心的动量分布，还要得到极化的结构函数。

存在两种“普通”的极化结构函数，同非极化的情况一样，它们都依赖于父核子的动量百分比。对实验者们来讲，最简单易行的就是纵向结构函数。该结构函数说明，探针粒子的运动方向同核子的自旋方向可能相同或相反，这两种情况中散射概率会存在差别。这两种散射截面的差别同纵向结构函数成正比。如果换一种方式，利用横向极化的靶——靶粒子自旋方向同粒子束运动方向垂直，实验者们就可以得到另一个分布，即核子的横向自旋

分布。

纵向结构函数可以用夸克模型清晰而简洁地解释由核子中所有夸克和反夸克引起的总纵向自旋贡献。所有动量百分比都要算进去，由于上夸克和下夸克所带的电荷不同，上夸克贡献的重要性是下夸克的4倍。结构函数描述了自旋同父质子相同和相反的夸克之间分布的差别。上夸克、下夸克、奇异夸克及三种相应反夸克的味涵盖了所有的基础，即使由三夸克组成的最简单的核子模型，只考虑上夸克和下夸克的自旋贡献就够了。

实验表明，最简单夸克模型计算出来的磁矩与实验值惊人地一致，但用来预言自旋时却出了问题。根据这一模型，自旋平行于反平行于父质子的上夸克之间的差别是4/3。而下夸克和奇异夸克相应的差值分别为－1/3和0。因此，总差别的数值是一个整自旋单位1。换句话说，核子自旋100%由3种价夸克引起，这同实验观测到的结果不一致。简言之，简单夸克模型太简单了，无法解释核子的自旋。

求和规则岌岌可危

纯粹依靠思考来得到自旋结构函数的形式，是理论物理学家们很乐意做的事情，但是他们做不到。他们真正做到的是写下两条求和规则，表示在累计所有可能动量百分比后自旋结构函数的性质。

其中一条规律是由比约肯创立的，它就像茫茫大海上伫立的灯塔一样指引着研究方向，这片未知的大海就是自旋结构物理学。比约肯求和规则用弱相互作用参数，表达了中子和质子的纵向自旋结构函数之间的差别。中子自旋结构函数同质子自旋结构函数通过某种方式相联系并不是一条重大新闻，回想一下，同位旋对称性已经将质子和中子联系在一起了。再回想一下中子能够通过弱衰变过程转变为质子，弱相互作用参数的出现也是自

然的。

比约肯在 1966 年整理出了他的求和规则，这比 QCD 理论诞生的时间还早。他用到了相当一般的方法，而不是详细的夸克模型或 QCD 计算。比约肯并没有为他的成果激动，他认为该结果“毫无价值”，因为没有任何实验能够检验它。在这一方面，他低估了实验物理学家们的睿智，他们最近已经证明了比约肯规则的正确性。这种检测过程需要用到中子自旋的数据，人们早在 1993 年就得到了第一批可用的数据，但真正精确的数据直到 1997 年才出现。

随着 QCD 理论的诞生，比约肯求和规则通过它的一个预言被世人所知，该规则现在的版本还包括与强耦合参数有关的修正。经 QCD 修正过的比约肯求和规则，在百分之几的精度内同现代质子和中子数据吻合，它说明了 QCD 理论的成功：与比约肯规则相悖就会导致严重的问题。

在 1974 年，加州理工学院的约翰·埃利斯（John Ellis）和麻省理工学院的罗伯特·加菲（Robert Jaffe）指出了他们使用比约肯规则的困难。检验比约肯规则需要质子和中子的数据，但是在那时，根本没有可能在几年之内得到中子数据。因此埃利斯和加菲开始寻找一套新的求和规则，这套规则能够单独用于质子或中子。他们得到的埃利斯-加菲求和规则同比约肯的规则很相似，都忽略了 QCD 理论以及所有夸克模型的思想，但为了继续前进，他们不得不加入一个新的假设：奇异夸克同反奇异夸克对核子自旋没有贡献。他们甚至都没有假设核子中不存在奇异夸克，只是认为它们的自旋相互抵消，因此净自旋贡献为零。考虑到核子中不应该存在大量奇异夸克，这一假定看起来并不是很不合理。

尽管埃利斯-加菲规律并不是 QCD 理论的结果，但最近加入了 QCD 修正之后，它仍然是一个很好的假设。这或许是一件好事，因为这个规则不成立。它得到的纵向自旋结构函数的值与实

验值不同。EMC 最初就是用这种矛盾掀起了自旋危机。用夸克的语言来讲，埃利斯-加菲规则表示，大约有 60%的核子自旋来自于夸克本身的自旋——这个值是实验值的 2 倍。

自旋都跑到哪里去了

尽管埃利斯-加菲规则在实验面前失效了，但它还没有差到被马上遗弃。当然，它也没优秀到可以推动某些基本理论的修正。我们最好把它同实验偏离的事实看做是一次学习的机会。

首先，将实验测定的纵向结构函数转换为有关夸克携带自旋多少的描述是一项很难处理的工作。它如此深奥，以至于 SLAC 和 SMC 的实验者们，在好几年的时间之内都在夸克携带核子总自旋百分比的问题上有分歧。直到 1997 年，误解才最终消除，人们从实际数据出发，通过分析和推理解决了这一问题。

其中最大的一个问题，是人们争论的主要内容，并且现在仍然困扰着人们。这个问题集中在怎样将所有可能动量范围的贡献加到一起。困难在于，对核子自旋有贡献的夸克主要是占核子动量百分比很低的夸克，但实验者们很难测量出这些贡献。实际上，要想得到一个可以与理论对比的答案，就需要外推动量范围两端的值。因为我们无法测量这两个端的值，尤其是在低动量端，该怎么计算纯粹看自己的信仰了。如果存在一些实验物理学家们没有观测到的东西，而理论物理学家们又没有加以考虑，得到的结果就有可能出错。这或许可以解释埃利斯-加菲规则为什么会失效。

但是，也很有可能外推成功，两者之间确实存在差别。那么，奇异夸克对自旋没有贡献的假设就有可能不正确。如果奇异夸克的贡献大约占质子自旋的 10%，并且同质子本身自旋的方向相反，埃利斯-加菲规则的预言与实验结果就可以勉强统一起来。但是用微扰 QCD 理论计算海夸克的产生时，却无法得到这种奇

异夸克极化的结论。除此之外，非极化深度非弹性散射实验还表明，奇异夸克携带的动量不到核子动量的3%，这就产生了一个问题——为什么奇异夸克对自旋的贡献会如此之大。

或者，也有可能奇异夸克对核子自旋的贡献很小，但其他一些因素产生的自旋占了主导作用。比如有可能是夸克的相对运动（换言之，也就是夸克的角动量）、胶子自旋或者胶子之间的相对运动（胶子的角动量），这些因素都没有在简单夸克模型中体现出来。最近，包括来自于弗吉尼亚州纽波特纽斯市杰弗逊实验室的安纳托利·拉杜施金（Anatoly Radyushkin）在内的一些理论物理学家，提出了所谓的推广部分子理论，试图涵盖核子中各部分子之间的作用。它能够帮助我们理解部分子相对运动对核子总自旋的贡献，甚至对夸克和胶子在核子中分布的影响，它就像是一种给核子拍的X光片。但是，我们也要付出一定的代价，那就是我们要设计更好的实验识别碰撞碎片。

HERMES就属于为此设计的新一代自旋结构实验。到此为止，自旋实验的目光都集中在测量散射的探针粒子上。但是经过特殊设计的HERMES实验能够探测和识别靶核子留下来的单个碎片。这些粒子——例如π子和K子，会包含被撞击的夸克，见图8.18。因此，包含奇异夸克的K子的数量就说明了靶核子夸克海中参与碰撞的奇异夸克数。用这种方法，实验者们就能够测出奇异夸克对核子的贡献。的确，HERMES的实验者们证明了海夸克的贡献很小，并且他们不支持奇异夸克自旋与父核子方向相反的想法。他们也没有观测到任何海夸克同价夸克贡献抵消的证据，而这是一些研究者解决自旋危机的一种想法。

实验者们还提出了一种方法，以获取胶子对核子自旋贡献的重要信息。理论物理学家们能够计算出，自旋结构对探针波长的依赖怎样反过来受到胶子自旋分布的影响。实验物理学家们已经开始研究这种关系，以获取胶子信息。而其他仍处于初期阶段的

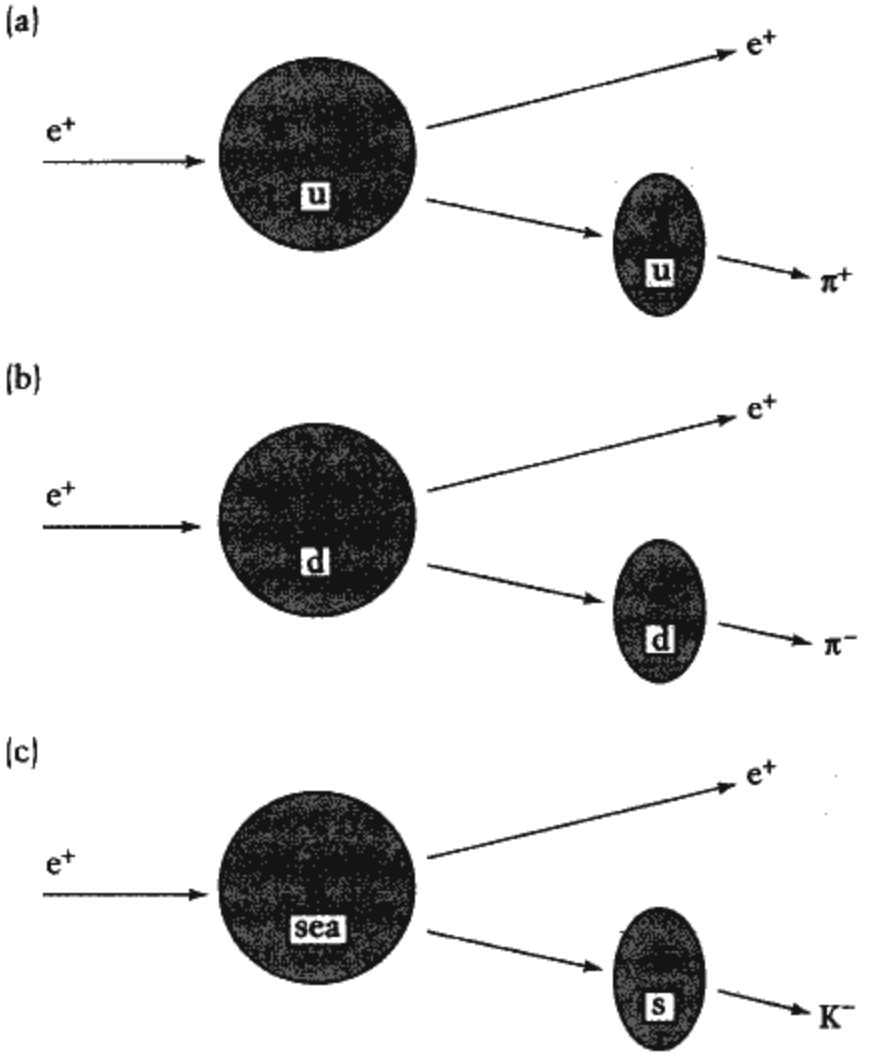


图 8.18　要识别 μ 子-质子或者正电子-质子碰撞过程中产生的 π 子和 K 子，实验者们可以测量不同夸克味道的贡献实现。出现在 π 子和 K 子中的额外夸克和反夸克来自于真空，这已经在第六章中讨论过了

一些实验正在努力测量“穷亲戚”——横向结构函数。按照简单夸克模型，这个函数也有简单而明了的解释：如果夸克是纵向极化的话，函数的值是零。SMC 在 1996 年完成了首次对横向结构函数的试验性测量，并得到了一个非零的结果，这表示核子内部

的夸克和胶子之间存在着自旋依赖关系，这也是高扭度效应的一个例子。HERMES的实验者们已经证实，胶子确实对核子自旋有贡献，它们的自旋方向同它们所在的核子相同。

简单夸克模型在处理重子磁矩问题时表现得相当出色，它还推断核子自旋100%由核子内部的三个价夸克引起。假设核子内部不存在奇异夸克极化，埃利斯-加菲求和规则用一般原理证明，总自旋的60%来自于夸克自旋。但实验得到的数值却为25%～30%。幸运的是，它比用QCD处理质子和中子自旋结构差别的理论——比约肯求和规则的预言要好得多。但尽管如此，那个大问题还是存在——核子的自旋到底来自何处？我们可能会将赌注押在夸克本身的自旋、胶子提供相当大的一部分自旋以及夸克相对运动的影响上。

坡密子：复活

其实没有必要指望用最新的自旋结构或夸克分布数据，来处理QCD理论面临的挑战。在QCD理论诞生之前，实验物理学家们已经开始搜集基本的质子-反质子散射信息了，这些信息现在仍然令QCD理论进退维谷，因为理论不得不完全满足实验的要求。人们已经找到了用雷杰理论，尤其是一种称作坡密子（pomeron）的东西，描述有关总散射截面和弹性散射截面的强相互作用数据的方法。一事成则事事成，这一成功使坡密子风光了30多年，在它能够处理的问题上，坡密子表现得都很出色。但传统的坡密子和QCD理论是一对不和谐的伙伴。尽管QCD可能是一种基本理论，但令人尴尬的是，在处理大多数高能质子-质子和质子-反质子散射问题时，坡密子都要优于QCD理论。

那么到底什么是坡密子呢？HERA得到的结果清楚地揭示了这种神秘的物质，但大多数物理学家都忌讳将其称为粒子。它或

许是一种粒子，或者至少是质子或其他强相互作用粒子的一种可辨认的态。它是一种和胶子有关的东西。不管它是什么，坡密子，或者某种同它等价的东西，在强相互作用中发挥着举足轻重的作用。

家族神话

总散射截面表现了相互作用的总强度。在低能情况下，发生强相互作用粒子的总散射截面由于存在共振，会出现峰值，在许多情况中，随着碰撞能量的增加，总散射截面的值会不断降低。但在较高的能量下，会出现一种新的特征，位于莫斯科以南 90 千米的索帕克夫质子同步加速器的实验者们，在 1971 年首次看到这一现象。在碰撞能量升高时，质子-质子、π 子-核子、K 子-核子以及其他碰撞试验中的散射截面都会有全面的上升。就好像粒子在高能状态下会膨胀，变得更不透明。

一套有关强力的理论应该能够解释这一基本性质。实际上，如果不能解释这种现象的话，至少也应该存在一种能表现它的简单理论，即雷杰理论——一种用交换介子来描述强子散射的方法。在 1992 年，两位英国物理学家——曼彻斯特大学的桑迪·唐纳齐（Sandy Donnachie）和剑桥大学的彼得·兰兹霍夫（Peter Landshoff），利用雷杰理论推导出一套简单的指数定律方程，用来完整地描述在很大的能量范围内强相互作用的总散射截面。因此，雷杰理论肯定不会一无是处。此外，尽管它不是一流的基本理论，并且它同 QCD 理论相处得也不太融洽，但它依然源自于一整套著名而合理的物理定律，因此也不可能被立即抛弃。

唐纳齐-兰兹霍夫结果中，有一部分对应着传统的雷杰介子家族方法。例如其中一个家族包括 ρ 介子、a_2 介子、ρ_3 介子和 a_4 介子，另一个家族包括 ω 介子、f_2 介子、ω_3 介子和 f_4 介子。这种交换机制能够解释在低能端，散射截面发生的典型下降现象。

散射截面在高能区域上升的现象可以用交换另一种相关物体来描述，杰弗里（Geoffrey Chew）和史蒂芬（Steven Frautschi）在1961年为此专门引入这一概念，它以苏联物理学家坡密朗丘克（Issak Pomeranchuk）的名字命名，叫坡密子。

坡密子交换是一种神秘的介子家族中的介子发生的交换。也就是说，坡密子是一系列满足雷杰理论要求的粒子，它们的性质不同于任何已知介子。在能量非常高的情况下，坡密子交换占统治地位，有大量的实验证明情况确实如此。作为一个基本机制的一种解释，坡密子并不能让人信服。但如果用坡密子对其进行描述，结果却相当好，在它们自己的地盘上，坡密子甚至比QCD理论还要出色。

坡密子现身

它们的地盘就是所谓的软相互作用，在该过程中，参与碰撞的粒子之间几乎不传递动量。软相互作用总散射截面很小，类似于弹性散射过程，在弹性散射过程中粒子会互相弹开。软相互作用会给QCD理论带来麻烦，因为它所涉及的能量太小了，以至于无法使用微扰论。但是有一类软相互作用能够体现出坡密子的存在以及它们的性质，至少也能告诉我们它同QCD理论的共同点。

在某些质子-反质子碰撞过程中，粒子受到的撞击可能会比弹性散射更加剧烈，这使粒子会像一口大钟一样响起来，然后发生衰变或碎裂。如果一个初始粒子没有受到任何损伤，只是动量稍微变小了一些，而另一个粒子却发生碎裂，这种碰撞称作衍射散射（diffraction scattering）①，有时也称作衍射分裂：质子-

① 译者注：也叫小角度散射。

（反）质子碰撞中有15%是衍射散射。在一次典型的质子-反质子衍射散射事件中，一个初始粒子——比如说一个反质子，会携带着初始动量的95%继续前进。在相反方向上，会出现一束粒子，它们是另一个粒子所产生的碎片，在该例子中是质子碎片。值得注意的是，这些碎片的方向分布于反应区域中沿粒子束方向的一个半球之内，大部分探测器都探测不到粒子①。

在1985年，加利福尼亚大学洛杉矶分校的彼得·施莱恩（Peter Schlein）和CERN的贡纳·英格尔曼（Gunnar Ingelman）推断，由于坡密子传递强相互作用，它们本身也应该是由胶子甚至夸克组成的。根据他们的理论，在上面提到的质子-反质子衍射散射过程中，反质子放出了一个坡密子，然后继续前进。质子同坡密子相撞，产生一束粒子，由于坡密子只带有其父粒子的很小一部分动量，因此粒子束飞向原质子束的方向。质子-坡密子碰撞本身可以被看做坡密子中的胶子同质子中的夸克或胶子发生完全碰撞，见图8.19。在衍射散射事件中，质子或反质子中的一个保持完整，因此总体上至少有一些衍射散射事件会显示出独特的喷注特征，成对的喷注会从坡密子-质子或坡密子-反质子“子碰撞”中产生。这同质子和反质子都被撞碎的喷注产生过程有很大差别。

CERN的UA8合作小组，寻找并最终找到了这种衍射散射喷注，他们在1988年首次得到了试验性的结果。随着数据的增加，施莱恩（Schlein）所在的UA8小组发现，喷注对是背靠背产生的，且具有QCD喷注的典型特征。但是他们却无法将整个过程划分为两部分——坡密子散射和坡密子组分的散射。他们也无法测量坡密子内部所包含物质的结构函数，也不能检验坡密子本身在质子内部有某种分布的想法。

---

① 译者注：因为该方向没有探测器。

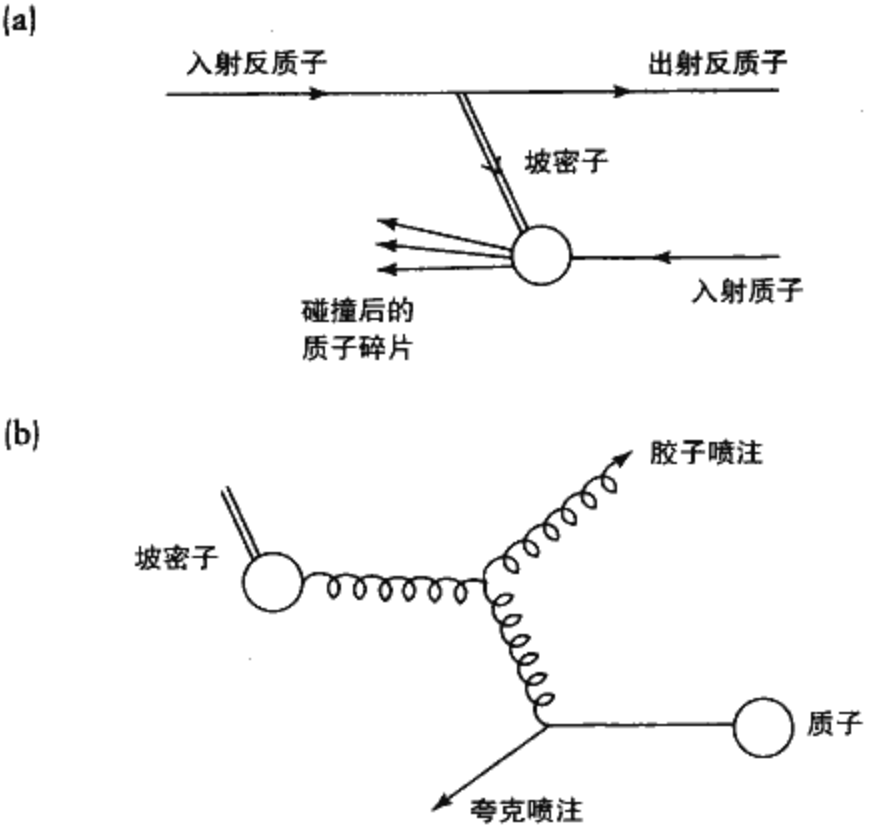


图 8.19 图 (a)：质子-反质子通过交换坡密子发生衍射散射过程。图 (b)：坡密子中的胶子同入射质子中的夸克发生硬的“子碰撞”散射过程

HERA 与 Tevatron 产生了冲突，CDF 小组和 D0 合作小组观测了 Tevatron 上质子-反质子碰撞的衍射事件，发现坡密子的确主要由胶子组成，但也测量到一定的夸克成分。此外，费米试验室的实验者们还找到了质子发射出坡密子，同反质子发射的另一个坡密子碰撞的过程。HERA 上的实验说明，衍射散射表示质子毫无损伤地继续前进，在坡密子图像中，它所发出的一个坡密子同入射正电子发生碰撞，见图 8.20。这会导致深度非弹性散射事件，此过程中质子不会碎裂——这些所谓的衍射深度非弹性散射事件由 H1 合作小组在 1995 年首次发现。他们的结果支持了把

整个散射过程分成两步的想法，首先是坡密子的发射，接着是坡密子内部点状组分的硬散射。此外，H1 合作小组以及 ZEUS 小组，还能测量质子内部坡密子的分布以及坡密子本身点状散射中心的分布。

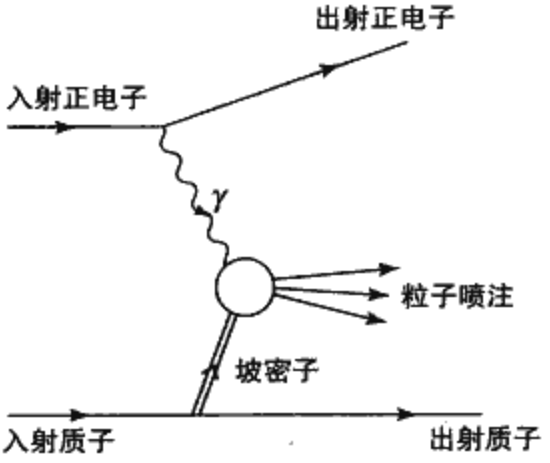


图 8.20　HERA 上的深度非弹衍射散射

尽管坡密子仍然高深莫测，但在当了 30 年描述散射数据的虚构工具之后，它已经显现出一些真实粒子的特征。一些人认为它是质子内部的一个态，或者一组态；另一些人则认为它是一种胶子的束缚态。或许它就是一个胶球——考虑物理理论的经济性，这当然是一种诱人的选择。另一些研究者认为，坡密子仅仅是描述过程的一种有用的术语，它的本质解释隐藏在 QCD 理论中，将来的某一天，我们也许会把这个术语扔进垃圾堆。在一些比较传统的人看来，坡密子仍然是雷杰理论的组成部分，在本质上属于非微扰范畴。这就是唐纳齐-兰兹霍夫坡密子或非微扰坡密子，也可以称作软坡密子，有时它被描绘成一对（非微扰）胶子。参与硬碰撞过程的坡密子吸引了微扰 QCD 理论学家，也导致很多人支持强坡密子的存在。强坡密子也叫做微扰坡密子或 BFKL 坡密子。QCD 理论的这个孩子可以利用 BFKL 方程并考虑一整套胶子的贡献来加以说明。

无论坡密子的真正本质是什么，有一件事情可以肯定：它可以成功地描述弹性散射和衍射散射过程，这些过程都是强相互作用物理的一部分。人们相信强力可以用 QCD 理论描述。因此，坡密子和 QCD 理论最终必定会以某种方式结合在一起。

第九章 庸人自扰

不招人喜欢的法国医生和讽刺作家弗朗索瓦·拉伯雷(François Rabelais)曾辛辣地写道："自然讨厌真空。"如果他所说的"真空"即传统意义上抽走空气、切断光源后的盒子内部，那么拉伯雷显然超越了他所在的时代：自然确实特别讨厌真空，讨厌到将真空完全填满。19世纪的科学家们曾认为，有一种物质充满了真空，这种虚构的物质叫以太，根据量子场论，这种物质实际上是存在的。

场论物理学家不讨厌真空，他们对真空进行了重新定义。在QED或QCD这样的量子场论中，通常所讲的真空指低能环境，当所有东西都切断之后它仍然存在，系统无法达到比它更低的层次。这就像是音乐中心的扩音器残留的嘶嘶声，即使音量表的指示为零，噪声也会持续。或者像是"空"咖啡杯底留下的残渣，抑或是"平静"的池塘表面上最后一个微小的涟漪。

曾有一段时间，人们认为真空态几乎不会产生任何物理效应，真空中只有少量残余的能量，量子力学称之为"零点能"。这仅仅是简单地规定能量图上能量零点应该在哪里，这样，相对于这个最低态的能量变化才会在实验中体现出来。但真空用了一种直接得令人吃惊的方式将自己表现出来。在最大的尺度上，真

空或许就是一种背景能量密度的源头，可以用一种叫做广义相对论宇宙常数的东西表示，因此它同宇宙的总体结构有一定关系。物理学家们现在认识到，在 QCD 理论中，真空是理论中至关重要的一个方面，是全局对称性中的一个重要角色和夸克禁闭的参与者，还能帮助我们理解由自由夸克和胶子组成的、称作夸克-胶子等离子体“气体”的这样一种新物质形态是如何产生的。但是要真正看到真空大显身手，最好是在电动力学中。

来自真空的作用力

在一个密闭的盒子中，将两块导体平板平行放置，然后把盒子中所有的空气、所有的热源、所有的震动源等全部移走，观测导体平板会发生什么现象。一般来说，应该没有任何现象发生：导体板位于真空中，不存在任何东西能导致某些现象发生。但实际上，这两块导体平板之间有一种吸引力，这种力量的起源就是导体平板周围虚无真空中残留的能量。

这并不是某种理论上的设想。实验物理学家们能够观测到平板间的吸引力，这叫做卡西米效应，是根据荷兰人亨德利克·卡西米（Hendrik Casimir）的名字命名的。尽管卡西米早在 1948 年就提出了他的解释，但直到 1997 年，当时在西雅图华盛顿大学工作的史蒂夫·拉蒙诺（Steve Lamoreaux）才在一次实验中，第一次真正令人信服地演示了该效应。拉蒙诺的实验最初只是一个学生的研究项目，实验装置由散落在实验室的仪器零件拼凑而成。从本质上讲，他用到的实验原理是一种扭转平衡。为了解决两个板平行的问题，他让一个弯曲的金属板同另一个平板相对，平板的一端连着一个可以自由转动的杆，杆的中点用金属丝悬挂起来。杆的另一端构成了一个电容器，可以通过在电容器上施加很小的电压确保平衡臂的稳定。这个实验十分灵敏，连他走进房

间时引起地板的倾斜也能记录下来，实验中，拉蒙诺将弯曲的金属板向连接杆的平板缓慢移动，并测量为了保持杆的稳定电容器上所需的外加电压。两板间的力的大小清楚地显示了对两板间距离的依赖关系，这是卡西米力的特征。

两平板之间的吸引力很小。对于两个相距 1 微米、表面积为 2 平方厘米的平板，它们之间卡西米力的大小大致相当于两个相距手指宽度的咖啡杯之间的万有引力，或者是一个氢原子中电子所受到的电吸引力。卡西米力来源于量子力学中的零点能，也就是那种永远不能被消除和切断的残余能量。在量子力学中，零点能可以被看做是海森堡不确定性原理的结果：如果一些量子物体能够完全静止地处于能量最低态，动量就不会有不确定性（它的值一定为零），不确定性原理就被破坏了。用量子场论的语言来讲，电磁场永远不可能被完全消除。电磁场的涨落永远不可能为零，表现为虚光子会在很短的时间内占据一定能量，它们会不断地产生，然后消失。那么这就是虚无，是能量最低的态，是真空的图像——一个沸腾的虚光子大锅炉。

正是这种活生生的以太，引起了激光中的“噪声”，也解释了拉蒙诺实验中金属板之间的作用力。以波动的观点来看，在两个导体板之间，光子可能的行动范围受到了限制，光子必须适应两板之间的空间大小。因此，不允许存在波长比两板之间距离的一半还要长的光子。但在板的外侧就不存在这种限制，波长较长的光子也能够产生和消亡。板外侧相对过剩的光子总体上就会形成一种净“辐射压”，将两块板向一起推。换句话说，将量子系统硬塞进一个密闭的空间内，会改变零点能，正是这种改变形成了可观测的金属板间作用力。换句话来说，真空并不是简单的一无所有，而是有很多起伏：在真空中放入物体——即使是比量子效应尺度大得多的物体，也能够以一种可测量的方式改变这种起伏。

卡西米效应的一个结论就是，在两板之间运动的光子，其速度比一般的光速还要大。但量子不确定性在发射或吸收光脉冲的时间上避免了超光速信息的传递，因此卡西米效应并没有给科幻小说迷们留下多少遐想的空间。在更普通的层次上，也会出现卡西米效应，比如戊烷可以溶解在水中，而其他类似化学物质只能在水中形成小液滴。另一个现象是超流体氦-4（^{4}He）很愿意爬上容器壁。卡西米效应还对胶体的性质有所贡献，胶体就是类似于墨水、黏土泥浆之类的系统，这些系统都是由小颗粒分布在液体中组成。20 世纪 40 年代，埃因霍温飞利浦实验室的荷兰人简·奥弗比克（Jan Overbeek）研究的正是这样一类系统——石英粉末的扩散，他发现流行的胶体理论存在缺陷。奥弗比克的修正引起了一位同事的兴趣，这位同事就是亨德利克·卡西米。卡西米认识到奥弗比克的修正是由真空的扰动引起的，这种效应现在以他的名字命名。

真空的作用在 QED 中很容易观察到，但它是量子场论的一种普遍性质。利用真空中的一些边界，例如导体板或者墨水颗粒，量子场的起伏就会发生改变。关键在于怎样观测到这些结果。

没有人进行过类似于史蒂夫·拉蒙诺双导体板实验的 QCD 实验来检测卡西米效应，但 QCD 理论的卡西米效应也真实存在。在麻省理工学院的口袋模型中就有这种效应的影子，这一模型已经作为一种核子模型在第七章中讨论过了。其中，系统的总能量中有百分之几来自于卡西米效应。但或许卡西米效应故事的真正寓意在于，真空具有重要的、意想不到的物理意义，我们必须认真对待。对于 QCD 理论，真空无处不在，它破坏了对称性，催生了奇特的新粒子，并且在夸克禁闭中暗示了自己的存在。

对称性、QCD以及虚无的空间

一小粒盐静静地躺在厨房的桌子上，这就是常识如何误导我们的一个例子。看起来静止不动的盐粒，实际上由于其内部原子的热激发在不断地嗡嗡振动。这非常像被量子化为光子的光波，这些晶格的振动由一些叫做声子（phonon）的小波包引起。在室温下，一小粒盐中会有大约10^{18}个这样的声子在嗡嗡振动。

声子的出现可以追溯到对称性上，或者可以说是对称性的缺乏：晶体是一种规则有序的点阵，在晶体中将原子结合在一起的相互作用，缺乏平滑的平移对称性或旋转对称性。粒子物理学家们最喜欢的例子——磁性也具有相似的特征。

就像在第五章中讨论过的那样，当铁从770℃以上冷却到该温度以下，会产生自发磁化。在一个磁畴中，一排一排单个的原子磁体指向完全相同，相邻粒子之间的顺排力最终战胜了温度引起的、使相对方向混乱的作用。对于一个指定的磁畴，可以明显观察到一种性质，这一性质意义深远：我们没有办法预言这些小原子磁体将民主选择出哪个方向作为它们一致指向的方向。我们可以这样考虑：相邻原子之间的作用力中不存在任何信息，描述相邻原子之间作用力的主方程中也不存在，但主方程中囊括了所有方向的所有信息。磁化系统从无穷多可能的方向中自发选择了一个。这就是被物理学家们称为自发对称性破缺的一个实例——能量最低态的对称性同涉及的相互作用对称性无关。

一些人认为，“自发对称性破缺”的表述方式有一点误导视听。磁铁处在能量最低的态——基态时，所有原子磁体的方向一致，它不具有那些原子磁体基本相互作用所具有的旋转对称性。因此，相互作用的对称性隐藏了起来，但并没有消失。他们认为隐藏对称性可能比自发破缺对称性的表达方式更贴切一些。但事

实上，还是后者取得了胜利。

自发对称性破缺现象在整个粒子物理中都能够体现出来，磁化现象会为我们说明一些原因。在基态上，铁磁体中的原子磁体都指向同一个方向。例如，如果所有的自旋都向上，那么磁铁中的一块水平切面有点像一片麦地，每一根麦子都代表一个原子磁体的箭头，麦穗就是箭头的头。夏日里的微风会在麦田中激起一片涟漪，与之类似，磁铁上的热量也会引起原子磁体指向方向上的涟漪，因此原子所在的平面上，方向箭头的顶端会发生移动，其运动方式类似于麦浪运动。

这些磁铁“箭头指向波”同光波量子化的情形极其相似，它们会产生一种无质量的粒子，叫做磁振子（magnon）。磁振子同“光波粒子”——光子类似，可以在中子散射实验中表现出来。

磁振子是自发对称性破缺的结果。在量子系统中对称性隐藏的情况十分普遍，在系统中，如果某种在理论主方程中可见的连续全局对称性没有出现在基态中，该系统一定包含声子和磁振子这样无质量、无自旋的粒子。像磁振子和声子这种粒子，都是由自发对称破缺引起的，它们统一被称为哥德斯通玻色子，以剑桥大学的物理学家杰弗里·哥德斯通（Jeffrey Goldstone）的名字命名，哥德斯通早在20世纪60年代就推断出有这种粒子存在。

在磁化的例子中，每一个磁化方向都可以选择，并且都具有同样低的能量，因此从理论上讲，从一个方向旋转到另一个方向不需要耗费任何能量。但原子磁体只选择了一个方向。这些本来能将一个基态转变为另一个基态的转动，却发现自己受到了限制，只能形成一些颤动，这些（量子化的）波动就是磁振子。因此，隐藏的相互作用转动对称性重新以磁振子源的身份出现。

自发对称性破缺以及无质量粒子也是QCD理论的一个特征。其中，能量的最低态是真空，在场论中，真空扮演着类似原子磁体系统“基态”的角色。而“无质量”的粒子其实是大家都很熟

悉的一位老朋友，它就是π子。

凭空产生的介子

情况是这样的。比较一下强相互作用特有的能量，3种轻夸克——上夸克、下夸克和奇异夸克的质量大致相同。就QCD理论而言，如果夸克质量相同，那么3种味道的夸克就可以随意交换，因为QCD理论并不关心夸克的味道。它表现为QCD理论的一种全局对称性，即交换3种夸克味道的对称性，它表明理论自动遵守关键的同位旋和奇异数守恒。

实际上在强相互作用中，典型的能量都特别高，将3种夸克的质量都当作是零也不会有太大影响。因此，可以将上面这种讨论再进一步，认为上夸克、下夸克和奇异夸克不仅质量完全相同，并且它们的质量都是零。这样会导致其他整体对称性的产生。

由此会引发一些新现象。因为这些夸克现在不具有质量，它们可以以光速传播。它们的自旋方向与运动方向相同或相反，分别对应着右旋或左旋的夸克。但在QCD中，夸克并不存在无质量夸克所具有的简单手性变换能力。一个左旋的无质量夸克可以发出一个胶子，但必须仍然保持左旋，右旋的夸克也会保持右旋。实际上，无质量的左旋和右旋夸克失去了相互对话的能力，它们只能够停留在各自的世界里。但是3种不同的味道仍然存在，现在左旋夸克和右旋夸克可以分别实现夸克味道混合，在数学上表示为存在味道转动。因此“无质量”表示，除了QCD理论中3个等质量夸克的整体对称性之外，还会出现一种双重对称性。

某些左旋物体产生的东西也必须保持左旋，而右旋的物体也只能产生右旋产物，这种变化称作手性（chiral），该名词来源于希腊语kheir，表示与手有关。开车去乡下就是一种手性变换，因

为如果不计一路上各种道路之间的“相互作用”，启程的时候如果靠路左（或右）边行驶，最后到达的时候仍然会在路左（或右）边行驶。粒子物理学家们将左旋夸克变为更多左旋夸克的对称性操作以及右旋世界中的相应变换，称为手性对称变换。

全局手性对称性看起来并没有给真实强相互作用对称性造成太大的影响。如果不是这样的话，每一个强子都会有一个与它质量一样的同伴，它们的各种其他性质也相同，但镜像对称性有差别。但物理学家们没有观测到这种“宇称同伴”。QCD 理论的手性对称被隐藏了起来，或者换句话讲，它发生了自发破缺。

对称性破缺的原因就是真空。真空不遵守理论的手性对称性，因此真空具有的对称性较少。在上面提到的铁磁性例子中，大量方向相同的原子磁体形成了一种能量最低的结构，它比磁体之间相互作用的对称性要少，因此产生了哥德斯通玻色子。手性 QCD 的情况与之类似，真空可以类比于能量最低的态，它不遵守手性对称，这特别像方向一致的原子磁体不屑于遵从相互作用的旋转对称性。

在铁磁性例子中，对称性破缺会导致磁振子。在 QCD 理论中，手性对称性破缺会产生一小批介子：3 种 π 介子（π^+、π^- 和 π^0）以及 4 种 K 子（K^0、$\overline{K}^0$、K^+ 和 K^-），还有最后一种也是最重要的一种粒子——η 粒子。如果对称性严格破缺，哥德斯通玻色子就不会有质量。如果对称性破缺并不严格，只是近似出现，结果就会产生低质量、无自旋的粒子，称作假哥德斯通玻色子，而不是无质量的哥德斯通玻色子。尽管真实的上夸克、下夸克和奇异夸克的质量与强相互作用能量相比很小，但它们并不严格为零，因此真实的 QCD 理论并不具有严格的手性对称，而只是近似满足手性对称性。因此手性对称性破缺会产生上面所列的 8 种小质量介子。通过这种机制，介子质量可以同（当前的）夸克质量相关。实际上，手性对称性破缺就是用来估计上夸克质量为 4

兆电子伏，而下夸克质量为 7 兆电子伏的那种方法，这些质量在第四章的表 4.6 中都列举了出来。能够通过结合形成像 π 介子、质子、中子等可观测粒子的夸克，其有效质量——换句话说也就是组分质量，大约是 300 兆电子伏，这是由夸克同周围手性对称性破缺真空相互作用的结果。自发对称性破缺以这种方式影响有效夸克的质量。QCD 理论的手性对称性特征也是另一种计算工具的出发点，该工具可以计算出大尺度上传统 QCD 微扰理论无法解决的问题。这种手性对称性破缺理论的出发点是一种完全手性对称性理论，即一种关于无质量夸克的虚构 QCD 理论，然后再将真实质量以微扰形式引入，由此得到的有效场论能够得到大量低能结果。以上提到的夸克质量估计就是一个实例。

那么 QCD 中的真空看起来是什么样子，这种“虚无空间”怎么会具有对称性破缺的性质呢？磁化现象再一次为我们提供了线索。其中的哥德斯通玻色子——磁振子就是磁化物中的涟漪，在磁体中沿某个任意方向的排列破坏了对称性。在 QCD 理论中，介子从对称性破缺中产生，这说明真空一定存在某种夸克-反夸克背景，物理学家们将其称为夸克凝聚。正是这种夸克凝聚真空中的涟漪形成了实际可观测的介子。

夸克凝聚是一种夸克-反夸克海洋，夸克对的自旋方向相反，以保证洛仑兹不变性和能量最低。因此，夸克凝聚建立在夸克同反夸克的相互关系之上，由于存在零点能，夸克和反夸克会不断地产生和湮灭。左旋夸克现在感觉到了大致相当于核子大小的空间范围内存在右旋夸克，正是它破坏了手性对称，因为左和右之间有了交流（图 9.1）。

实际上，夸克凝聚仅仅是故事的一部分。真空还是胶子凝聚以及其他一些组分居住的地方。背景夸克和胶子正是通过这种凝聚进入了 QCD 理论的求和规律方法，它对于计算强子性质十分有用，这在第七章已经介绍过了。但是 QCD 理论的真空结构，

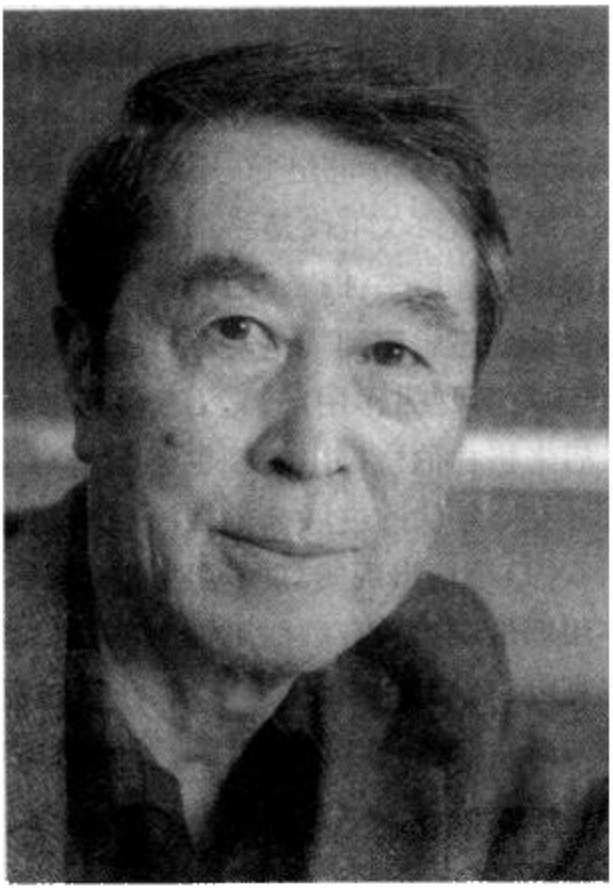


图 9.1 南部阳一郎，摄于 1994 年，他是第一批意识到真空和自发对称性破缺重要意义的人之一，对粒子物理作出了许多重要贡献。他的另一项贡献是同韩武荣与格林伯格各自独立提出了颜色量子数。图片来源：芝加哥大学

仍然存留着一个巨大的尚未解决的疑团。尽管理论物理学家们都坚信，真空在夸克禁闭中起了重要作用，但它仍然令人迷茫。用真空的语言来说，夸克禁闭的双超导模型——也已经在第七章讨论过——认为，一个 π 子就是一对由色-电场管道连接起来的夸克和反夸克，它生活在真空中，而真空中聚集着色-磁单极和反单极，是一种色-磁超导体。当理论物理学家们真正理解夸克禁闭的时候，他们才能揭开真空之谜，至少部分上是这样。或者他们必须先理解真空，然后才会理解夸克禁闭。不管是哪一种方

式，没有人敢保证问题会迅速解决。理论物理学家们在 QCD 理论的曙光来临之前，就已经考虑夸克凝聚和对称性破缺的问题了，它源自于马文·戈德伯格（Marvin Goldberger）和萨姆·特雷曼（Sam Treiman）在 20 世纪 50 年代研究 π 介子衰变时所做的基础性工作。20 世纪 60 年代是一个神秘肆虐的时代，当时存在着哥德斯通的定理，后来彼得·西格斯（Peter Higgs）和其他人将自发对称性破缺思想用到规范对称性上，这最终对电弱理论的诞生起到了至关重要的作用，当然还有盖尔曼和尼曼的八正法、夸克以及后来格林伯格提出的颜色。但是对于强相互作用而言，整个 20 世纪 60 年代的进展都是在没有一套强力理论的基础上做出的。到 QCD 理论出现的时候，理论物理学家们已经知道，一套强相互作用理论应当具有近似手性对称性，并且这种对称性会由于存在夸克-反夸克凝聚而自发破缺，该过程与 π 介子的质量和寿命都有关系。QCD 理论能够自动容纳所需的对称性模式，这也是它能够被迅速认可的一个原因。

对称性审计

尽管手性对称性及其破缺取得了大量成功，但实际上还隐藏了一个问题。为对称性算一笔小账就可以发现错误的地方。如上所述，等质量的上夸克、下夸克和奇异夸克赋予 QCD 主方程一种新对称性。这种对称性即 U（3）对称性，如果忽略技术细节，它是我们的两位老朋友 SU（3）群和 U（1）群的乘积。假定夸克没有质量，就会引入另一套 U（3）群，换言之就是另一套 SU（3）群和 U（1）群的乘积。其中，一套 SU（3）群对应八正法，它将发生强相互作用的粒子都集中到各个家庭中。π 子和这些家庭的其他介子，与第二套 SU（3）群的自发对称性破缺有关。一套 U（1）群用来保证重子数守恒。因此一切看起来还令人满意，除了剩下的一个 U（1）群，它似乎找不到可观测粒子

的对应。或许它是另一种对称性破缺的牺牲品，是这样么？显然不是——并不存在相应的轻介子对称性破缺。整理一下表面上看来是技术细节的问题，却给理论物理学家们带来了这么大的麻烦，他们给这个问题起了个名字，叫做“U（1）问题”。温伯格第一个注意到U（1）问题的重要性，并确保它列在理论物理学需要解决问题的名单里。令人吃惊的是，U（1）问题的解决也许涉及真空结构的另一个令人吃惊的层次。

在真空中迂回前进

大家可能会想，QCD理论的研究已经那么深入了，对称性——或对称性破缺的每一个小部分都应该已经研究过，并且都得到了解决。但事实不是这样。最后一种U（1）对称性现在仍然没有驯服，它预示着还存在许多在现实世界中不存在的粒子，所有像质子这样的普通粒子都有对应的粒子，它们与普通粒子的各种性质完全相同，只是内禀镜像对称性相反。这种宇称对应粒子并不存在，这说明U（1）对称性通过某种方式发生了破缺，问题在于它是怎样发生破缺的，因为并没有适当的介子可以标记这种自发对称性破缺。该问题的答案揭示出，QCD理论的真空中蕴藏着出乎意料的丰富内容。一个整理小花园的例子会告诉我们究竟是怎么回事。

想象一个风和日丽、阳光明媚的周末。在郊外，池塘旁边摆放着夏敦埃酒，蜜蜂绕着花丛嘤嘤飞舞，鸟儿在树上歌唱。某个人像往常一样用除草机打破了这片宁静。马达除草机最不好，因为它的噪声很大。电动除草机比较安静，但是园丁却会遇到一个无法避免的问题：电源线问题。

给一块长着两棵树，还有一个温室以及一些其他障碍物的草坪除草，用电动除草机就会有些麻烦，因为电源线会在障碍物之

间来回缠绕。这最终会限制除草机的移动——我们无法推着除草机穿过树干，也没有办法把电线从树顶上绕过去。理清电线的唯一方法就是沿着原来的路径返回去。当然，就算绕树走大圈也无济于事，电线还是会缠绕，从这种意义上来讲，短路径和长路径完全一样。绕同一棵树沿同一方向转两圈会使事态更加严重，同只绕一圈的情况明显不同。

所有这一切在任何一位园丁看来都是很明显的事实，因此他们本能地掌握了同伦（homotopy）理论，这种伟大的数学分支思想之一被称作拓扑学（Topology）。拓扑学涉及对空间进行分类，而不关心两点之间的实际距离。拓扑学家们耗费了大量时间和精力研究被他们称作“拓扑不变量”的东西，他们可以用它来区分两个空间在拓扑意义上是否相同。他们并不是按照常识描述事物——在一个拓扑学家眼中，足球同橄榄球完全一样，茶杯同（环状的）油炸甜饼圈也完全相同。一位爱好拓扑学的园丁可能会绕着树走一圈，然后把这一圈标记为“回转数”1；绕着大树转两圈之后，回转数就变为 2，等等。从拓扑学的观点看，绕大树走的路径，无论长短都是一样的，因为两者产生的回转数都是 2。稍微正式一点讲，如果两个路径能够通过连续地变形而互相转化，转化过程中路径不截断，那么它们就完全相同。绕一棵大树旋转两圈同绕一棵大树转一圈的路径不同，并且它们都与绕两棵大树转圈的路径不同，因为存在障碍物——在该例子中障碍物是大树，这几种路径无法相互转变。

推着除草机绕大树走一个圆圈回到出发的起点，看起来就好像什么也没有发生过，但初态和末态还是有一点细微的差别，因为电线绕过障碍物产生了差别。

将一个电子分割，让它部分通过电流线圈的一侧，部分通过另一侧，然后在远处复合为完整的电子，结果就会形成一个环绕障碍的封闭电子路径。当一个电子波被分割然后重新复合，结果

会得到一种干涉条纹。当它被分割，在复合之前通过螺线管的两边时，干涉条纹会发生改变，这就是我们在第三章曾遇到过的阿哈罗诺夫-玻姆效应。它完全等效于简单地让电子绕螺线管转动的情形。在这种情形中，起点和终点在物理上可能有所不同，但电子的改变只依赖于封闭在螺线管内的磁场，而与路径或距螺线管的距离无关。除草机对草地的非平凡拓扑性质十分敏感，因为草地是一个布满大树和其他障碍物的平面。阿哈罗诺夫-玻姆效应显示了无磁场真空的非平凡拓扑性质，它是一个散布着小孔的平面，这些小孔即螺线管。

阿哈罗诺夫-玻姆效应表明具有非平凡拓扑性质的真空怎样产生物理上可观测的结果。在 QCD 理论中也存在非平凡真空拓扑性，它也会产生物理效应。

利用数学家们的技巧，在 QCD 理论主方程中加入一个全新的部分是一件很简单的事情，这一部分揭示了 QCD 真空的拓扑属性。这里所谓的拓扑术语就是规范和洛仑兹不变性，它们满足理论最基本的对称性，但这仍是比较新奇和不寻常的，因为它们对任何费曼图都没有贡献，也同已确立的、粒子动物园中的粒子家族没有任何关系。

1975 年，亚历山大·伯利亚科夫（Alexander Polyakov）同他在莫斯科朗道理论物理研究所的同事一道，首先说明了这一额外的拓扑部分是怎样扩展物理概念的。对于 QCD 理论，他们的工作揭示了一种胶子实体，该实体从很久以前很远的地方诞生，并在将来且很远的地方终结，但在途中需要一个规范变换“纠缠”。尽管在最初和最后，场强都会消失，但中间存在一个局域区域，其中的场具有一定的正能量。伯利亚科夫和他的同事发掘的物种现在被称作瞬子（instanton），是由胶子形成的一种“能量块”，是胶子场中的一种起伏。但它是一种十分特殊的涟漪。

事情是这样的，在 1834 年，工程师约翰·斯科特·罗素

(John Scott Russell) 在爱丁堡-格拉斯哥运河旁骑马，他看到一条由两匹马拉着的小船。船突然停了下来，在船头前面形成了一个很大的圆形水波。罗素对这个现象很感兴趣，就骑着马跟随着这个水波走了两三千米，直到看不到水波才停下。在随后向英国皇家联合会描述这种奇特的孤立波动时，罗素对现在被称为孤波(soliton) 的东西进行了首次介绍。

顾名思义，孤波就是一列单独的波动。它与将石头扔到池塘里形成的那种普通波动不同，那种波动会很快扩散并且消失，但孤波能够保持其形状和大小。当两列孤波相撞，它们能够相互通过，在另一端再次呈现其原来的形状。在 20 世纪 60 年代，物理学家们和数学家们都开始关注孤波现象，孤波为物理学家们提供了一种基本粒子模型的自然方案，而数学家们也发现它们是装满数学宝藏的阿拉丁山洞。

我们已经遇到过一个孤波的例子。特霍夫特和伯利亚科夫(图 9.2) 在 1974 年提出的磁单极子就是一种孤子，它是瞬子的一个表亲。瞬子是孤子的另一个实例。作为一种孤子，瞬子能够

图 9.2　亚历山大・伯利亚科夫 1978 年在美国举行关于瞬子的首次演讲。伯利亚科夫是第一批认识到拓扑思想在杨-米尔斯理论中具有重要作用的人之一。图片来源：亚历山大・伯利亚科夫

四处闯荡而不会消失，也能够互相穿过且毫发无损。它们非凡的稳定性来自于它们所具有的拓扑性质。它们携带着一个量——“拓扑荷”，这个量是守恒的，不会轻易被破坏。正是这种拓扑属性保证了它们的安全。这有点像两头固定在桌子上的一条纸带，但是在固定之前将纸带反转了180°。纸带打卷的地方很容易局限在某一地方，尽管这种卷可以沿着纸带来回移动，但除非纸带的一端松开或者纸带断裂，否则上面的卷永远不会消失。这种卷就是这个例子中的拓扑性质，它担任了拓扑荷的角色。瞬子就是能量在空间和时间上都局域化的胶子孤波——它被打了一种结，或者说被赋予了一种拓扑性质，这可以通过守恒的拓扑荷测量出来。

通常，守恒荷都联系着一种连续对称性，可以用主方程在某种对称群作用下保持不变来表示。一般，随时间变化的荷会产生流。例如，在麦克斯韦的电磁学当中，某个封闭区域内电荷减少的量，要等于流出该区域边界总电流的量。在经典理论中，可以将主方程进行路径变分，从而找到能量最小的路径，这样就可以推导出麦克斯韦方程组，这是最小作用量原理的一个例子。

但大体上讲，改变路径这类变换对系统的拓扑性质不产生任何影响。只要涉及拓扑性质，所有相互之间有“一点儿差别”的路径都完全等价，就像除草机的电源线紧紧缠绕着大树和松散地绕在树上没什么两样，有意义的只是缠绕的圈数。

加在主方程上产生瞬子的额外项，能够表现出这种路径不敏感性，另外，它还缺乏通常荷和流表现出来的连续对称性。但存在一种守恒量，它源自于两个空间重叠的方式，除草机电源线的问题实际上提供了一个很恰当的原型。在除草机电源线的例子中，一条直电源线可以同圆形导线反复重叠，如果圆形导线包含一棵大树，每增加一圈就表示会产生新的效应。例如一个空间的圈重叠在另一个空间上——其中一个空间表示颜色对称群，而另

一个空间则是 QCD 理论起作用的宇宙——会变为编码在胶子波上的无穷多可能的套圈。“圈数”的精确数学表述是回转数，它就是拓扑荷。

这种新守恒荷会带来一种新的贡献，它对应量子理论的一种新结构。它来源于费曼路径积分试图包括所有结构中的所有贡献。因此我们有充分理由相信，路径积分会受到拓扑项的影响。实际就是这样。

在量子力学中，系统最开始可能处于能量最低的态——系统的基态，然后会经历一些跃迁或者其他过程，最终又会回到基态。在量子场论中，起点和终点都是真空，费曼的路径积分正是通过真空找到了正确的表达方式。在第三章我们引入路径积分时，目的是用它来计算从一种场量子组态经由所有干涉路径跃迁到另一种组态的可能性。也可以说，路径积分就是从真空到真空跃迁的振幅，表示为将所有连接初态和终态的路径都加起来。换言之，根据量子场论，一个过程的起点是真空态，场量子从真空中诞生，然后到一些地方干一些事情，最终灭亡。真可谓百川至海，落叶归根，从哪里来就要回到哪里去。路径积分可以计算出发生这种过程的振幅，计算要包括所有对主方程有贡献的路径，该振幅的平方等于发生跃迁的概率。

这当中出现的新内容是寻找包含扭结的路径，也就是那些具有拓扑差别的路径，这些路径不能连续地互相转变。因此瞬子对应着一条连接初末真空态，并具有不同回转数的路径。也就是说，瞬子为系统提供了一个通过某种方式相对于初始点实现“纠缠”的路径。由于联系真空的瞬子具有不同的回转数，“这种”真空实际上由无穷多看起来相同但拓扑性质不同的真空态组成，每个真空态都用不同的回转数标记。

存在许多拓扑意义上不同的态的想法，类似于无限长的除草机电源线反复在大树上绕来绕去。每一种纠缠状态都同缠绕圈数

更少或更多的态不同。可以投机取巧地将一种纠缠态变为另一种——通过在大树顶上绕圈来实现。对于真空也存在类似的壁垒：每一种具有拓扑差别的真空态都会被能量势垒同其他态分隔开。现在，量子力学中的势垒与经典世界中的障碍物并不相同。一个在足够深的经典篮子里的经典青蛙肯定无法逃脱，但是在篮子中的量子青蛙能够渗出篮子的边缘。这种发生渗漏的效应称为量子隧道效应，它在放射性 α 衰变过程以及隧道电子显微镜中都起着重要作用。瞬子通过将一个真空态同另一个具有不同拓扑性质的真空态联系起来，产生了这种泄露——量子隧道效应。QCD 理论中真正的真空，实际上由所有具有相同能量但不同回转数的可能真空态混合而成，每个真空态对真实真空的相对贡献由参数 θ 控制。QCD 理论中的真空实际上是这种无穷纠缠的沼泽，它被称为 θ 真空。

θ 参数是 QCD 理论中包含的另一个参数。它决定进入 QCD 理论主方程拓扑项的大小。它的值很小，实验物理学家们只知道这么多。理论物理学家们明白，由于马上就会出现的原因，如果不是这样的话，他们马上就会遇到另一个灾难。必须有一些特别充分的理由来说明为什么粒子物理学家们要将这样一种可能很危险的东西作为拓扑项引入他们的游戏。理由的确存在。

能量块拯救 QCD 理论

有时，对称性是量子的牺牲品。也就是说，有时出现在理论经典版本中的对称性，经受不起量子化的考验。实际上这就是在 QCD 理论当中，剩下的那种 U（1）对称性的命运。其中的关键就是要找到对称性破坏的源头。

关键如下。如果这种 U（1）对称性遭到了与 π 子和 K 子相关的普通对称性破坏，就应当存在一种相应的介子。为了适应该模型，这种粒子不仅应该具有真实世界中 η 介子的性质，还应当

具有同π子差不多的质量，因为两者具有共同的出身。但是，η介子已经被标记为一种哥德斯通玻色子，下一种具有正确性质的候选粒子η′粒子应该具有特别大的质量。看起来大自然并没有提供一种合适的粒子，来满足这种明显遭到破坏的对称性的自发对称破缺。解决这种难题的一种方法就是使用一种新的对称性破缺机制——一种能立刻破坏对称性的东西，但它并不要求存在一种类似π子的伴随介子。这样，U（1）问题就可以迎刃而解了。

瞬子能够提供一种合适的对称性破坏。一个无质量的胶子漩涡——瞬子，能将一个右旋的夸克反转为左旋夸克。这种夸克手性的反转就是手性守恒（或者称手性对称性）的破缺。因此，瞬子破坏了这种对称性，但又不要求存在一种轻质量的介子。U（1）问题解决了，QCD理论存活下来，继续为明天拼搏。

QCD主方程拓扑贡献的振幅——也就是包含瞬子的那一部分，要受到θ参数的控制。因此，理论物理学家们有时用θ参数描述对称性破缺就不足为奇了。θ参数选择了一种独特的真空混合方式，以产生真空态。从某种意义上讲，θ参数在真空中选择了一个方向，并且是在无限多可能的候选方向中选择了一个，这是一种对称性破缺。与之非常类似的是，一头大象呆在一根（非常结实但很柔软）的旗杆上，会使旗杆弯向某个特定的方向，这样就破坏了旗杆最初的360°旋转对称性。瞬子作为一种θ参数的起伏出现，这特别像在磁化系统的例子中，原子磁体指向在方向上的起伏导致了磁振子的产生。差别在于θ参数的起伏包含了一系列不同的可能真空态，这些真空在所有可能的方面都不相同。这是因为，这些起伏的产生者并不在乎QCD理论的基本颜色对称性。最终的总体结果就是，瞬子是非物理的。瞬子并不是像电子和π子那样的真实粒子。

杰拉德·特霍夫特发现了瞬子在U（1）问题中扮演的角色，罗曼·捷卡维（Roman Jackiw）、克劳德·拉比（Claudio Re-

bbi)、柯蒂斯·加伦（Curtis Callan)、罗杰·戴森（Roger Dashen）和戴维·格罗斯等人的努力进一步提供了真空结构的详细内容，这掀起了一阵研究该问题的热潮。但随着理论物理学家们认识到瞬子不会轻易地泄露它们的秘密，这种热度很快就退却了。在特霍夫特完成了他的奠基工作 20 年后，温伯格写道："尽管最初抱有很大的希望，但瞬子的发现并没有给我们在 QCD 理论中进行定量计算带来多大的帮助。""另一方面……他对我们定性地理解 QCD 理论以及其他规范理论，带来了难以置信的巨大变化。"

瞬子并不是真实的粒子，这个事实并不一定表示它们无法观测。实验物理学家们在深度非弹性散射实验中，有机会通过利用瞬子能够反转夸克手性的性质，观测到它们的蛛丝马迹。一个电子同一个质子发生碰撞，可能会得到一种非常特殊的夸克混合物——电子发出的胶子可能同质子中的胶子散射，得到一堆胶子以及（比如是）右旋的夸克和反夸克，见图 9.3。在计算机模拟瞬子效应的协助下，HERA 的粒子物理学家们支持了这一反应的存在，尽管他们尚没有完成可用而明确的瞬子实验。他们认为，一个瞬子很可能会产生一个火球，生成 10 个左右的夸克和胶子，此外还有一束夸克喷注。火球的碎片很可能会比一次"普通"散射事件产生更多射向旁边的能量和更多带电粒子，我们要想办法了解怎样处理数据，以揭示隐藏在大量事件中的瞬子事件。HERA 还是一个研究极化质子与极化轻子碰撞的好地方。物理学家们认为，由于夸克同瞬子发生相互作用时会产生手性反转，瞬子对依赖于自旋的散射截面贡献应当与熟知的夸克-胶子碰撞不同。无论它的路径是什么，发现瞬子就能够将梦想变为现实——瞬子是一种全新的物体，一种具有拓扑性质的物体，也是真实存在的 QCD 内在真空结构的指示器。

实验物理学家们除了要寻找瞬子之外，还要进行一些其他方

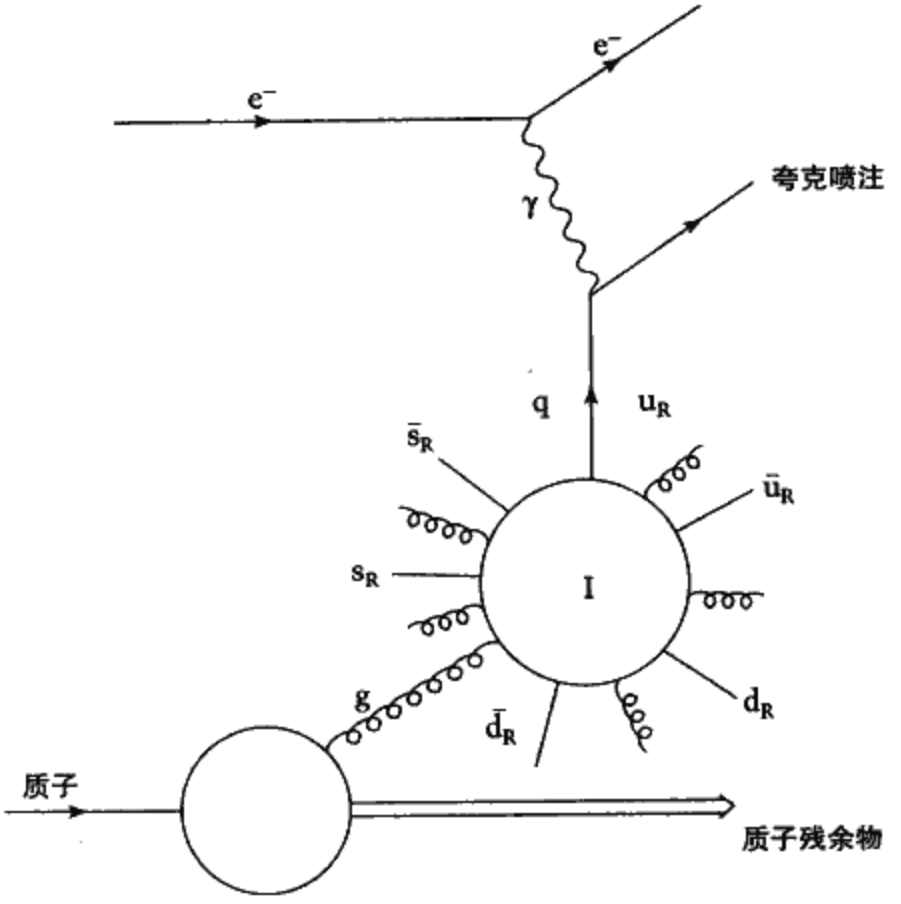


图9.3 在HERA上，假定瞬子（I）作用于电子-质子深度非弹性散射实验的示意图

面的工作——测量QCD理论中的θ参数。θ参数不仅在真空结构中十分重要，它还是令人恐怖的幽灵可能通过的一条密道。因为θ参数控制着理论中以拓扑为基础的那一部分，并且那一部分具有令人十分讨厌的性质，它既没有镜像对称性，也没有时间反演对称性。但人们认为，强相互作用遵守这两种对称性。因此，理论物理学家们面临着窘境：承认令人激动的真空拓扑性质能够通过瞬子解决U（1）问题，但这样会威胁到神圣的对称性。

实验者们获取θ参数值的方法是通过测量一种叫做中子电偶极矩的量，该量表示电场中的中子发生偏转的趋势。重要之处在

于，这个量的测量值能够为 θ 参数值的大小提供一个上限，该上限大约为 10^{-9}。换言之，θ 参数特别小，甚至有可能是 0，所有实验者都敢断定，它的值比十亿分之一还要小。理论物理学家们最初猜想的 θ 参数值大约是 1：理解 θ 参数值为何如此小的原因是一个全新的问题，业内人士将其称为“强 CP 问题”。

或许 θ 参数的值就是 0。人为地把它设置为 0 是一种欺骗手段，会导致其他复杂问题的出现。斯坦福大学的罗伯托·帕西（Roberto Peccei）和海伦·奎因（Helen Quinn）建议使用另一轮对称性破缺，用来解释 θ 参数为什么如此之小。他们的想法被温伯格和维尔茨克各自采纳，这又导致了另一种粒子的引入，一种物理的、无自旋的粒子，叫做轴子（axion），这是维尔茨克根据一种家用清洁剂的牌子起的名字。理论物理学家们最初设想的轴子同实验结果不一致，他们最终决定轴子应当具有很小的质量，或许只有几微电子伏，这说明实验物理学家们在寻找轴子的时候找错了地方。他们还估计，我们周围的轴子数量非常大，它们或许以每立方厘米有 1 万亿个轴子的密度充斥着宇宙。这幅图像立刻吸引了宇宙学家们，他们将轴子作为他们称为暗物质——宇宙中丢失的那部分质量——的候选者之一。星系在星系团中运动，表现出来的质量比天文学家们观测到的大很多，这只是大量的证据之一，证据表明宇宙中有大约 90%的物质无法观测到，因为它们不能发光。根据宇宙学家们的结论，暗物质的真实性不容置疑，但没有人知道暗物质到底是什么。它或许是很普通的东西，比如大量像行星或死掉的恒星这种很小的昏暗物体，或许是数量众多的中微子，也很有可能是轴子。

1983 年，佛罗里达大学的皮埃尔·斯奇维（Peirre Sikivie）指出，轴子在强磁场中很可能会转变为光子。这开启了新一代寻找轴子实验的大门。位于加利福尼亚的洛仑兹利弗莫尔国家实验室就进行了一项此类实验，该实验由莱斯利·卢森伯格（Leslie

Rosenberg）和卡尔·范·比博（Karl van Bibber）领导，实验的基本原理是用一个巨大的超冷却状态的射频接收器，去寻找由具有特定质量轴子产生的射频光子，见图 9.4。由京都大学的松木正史（Seishi Matsuki）领导的另一项实验，同样也依赖于轴子转变为光子的过程，但该实验利用原子束技术探测这些光子。CERN 的太阳轴子望远镜（CERN Solar Axion Telescope，简称 CAST）是一种新式的实验仪器，它用一台 10 米长的磁体（从 LHC 上退役下来的测试磁体）对太阳进行观测。太阳发出的轴子会在磁体内部转变为 X 射线，X 射线能够在仪器的另一端探测到。到目前为止，尽管人，们已经进行了大量实验，但还没有人发现轴子。根据一些研究者的想法，他们永远也不可能观测到轴子。这当然也给搜寻工作增加了新的理由。

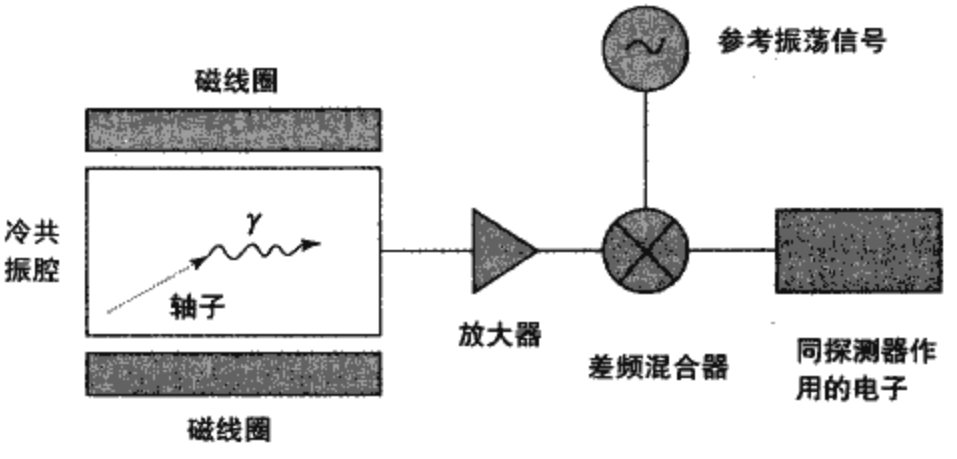


图 9.4　俘获轴子的一种常规方法。在施加强磁场之后，轴子会在共振腔内转变为光子。放大后的高频信号混合着参考信号，可以产生能被探测器接收的脉冲

真空的尾声

从以上众多的内容当中，可以提炼出两条简单的信息。首

先，真空并不是一种死气沉沉的虚无状态，它具有某种结构，并且能够产生可观测到的效应。其次，QCD 理论中的真空可以通过某种方式纠缠、打结，普通的微扰方法无法解决这种现象，它代表了理论的一整套更根本的结构，同夸克禁闭、瞬子及其他内容都紧密相关。真空远比人们平时想象的复杂。

第十章 棋盘 QCD 理论

QCD 理论中有一座高山，人们会发现，迟早，这座山将挡住几乎每一条道路。问题很简单：就是怎样在低能条件下处理 QCD 问题。

这就是在 QCD 理论中相互作用强度变化反映出的问题。正如在第五章讨论的，在高能和短程条件下，耦合参数相对较小，可以用传统的微扰论解决问题。但在低能情况下，也就是大尺度条件下——大约 1 飞米的量级上，QCD 耦合参数的值不再足够小，微扰论就会失效。为了推翻这座大山，理论物理学家们需要一些不用微扰理论，就能进行 QCD 计算的方法。这种工具有一天也许能开辟证明夸克禁闭的道路，或许能够揭示 QCD 真空的本质。理论物理学家们将能够用他们基本的夸克砖块，计算质子和其他复合粒子的性质，那样就能够严格计算出夸克和胶子碎片怎样形成了实验者们在强大的探测器中观测到的粒子喷注。像这样的可能性，我们可以列举出很多很多。

坏消息是，在总体上还不存在一种系统的数学方法能够做到这一点。但好消息是，利用人们建造的一些最强大的计算机，可能会解决一些问题。利用 QCD 理论计算，而不依赖于传统的微扰理论，这样一条通往非微扰 QCD 计算之路可能是很漫长的。

在计算机上模拟 QCD 的方法称作格点 QCD，它在量化低能、长距离下强力的性质方面具有巨大的潜力。但使用计算机进行计算的成本十分高昂，还会耗费大量的时间。这些额外的限制现在显得越来越严重。

计算机上的 QCD

肯尼思·威尔逊（Kenneth Wilson）在 1974 年引入了格点 QCD 的重要思想，它看起来十分简单。尽管时空是连续的，威尔逊的做法就是将它们划分成小块，用一些点和格子来表示。用一些钟表来划分时间，钟面上的秒针嘀嗒嘀嗒地从一个时间点指向另一个时间点。分布在河流上的踏脚石，也对空间有类似的效果。把时间和空间表示为格子上的点，然后让格点之间的空间完全消失，同时令整个格点阵的体积膨胀到无穷大，这样我们就回到了真实的连续世界。

将 QCD 理论应用到格点上，能够自动限制困扰理论的无穷大问题，并且能确保“普通”微扰论方法重正化。这是因为，如果一种晶格能够恰当地描述它所承载的夸克，夸克的波长（回想一下，夸克也是一种物质波）必须至少为格子间距的 2 倍。因此有限的格子间距能够提供长度尺度上的低端限制，这非常有用，因为短的长度尺度是（紫外发散）的源头。基本策略就是，先写下某有限晶格间距及有限总晶格体积情况下的 QCD 方程，然后利用大型计算机上的数值模拟，进行一些我们感兴趣的非微扰计算，最后通过令晶格间距缩小为 0 以及晶格体积趋于无穷大，让计算恢复到真实。这听起来很简单。

分布着夸克和胶子的格点有四个维度，一个时间维度和三个空间维度，它并不是一种物理上存在的格子，而是计算机中的一组数字。不过晶格可以描绘成一张由相交的线组成的网，其二维

示意图见图 10.1。两条线相交形成一个格点，在每点上有一个表示在该点存在夸克概率的数字。两个格点之间的连线将这些点连起来，胶子就位于这些连线上。图 10.1 显示了两个特别简单的结构。其中一个是一条线连接着两个夸克，另一个是没有连接任何夸克的圈，这个圈表示一个只有胶子系统的胶子自作用。

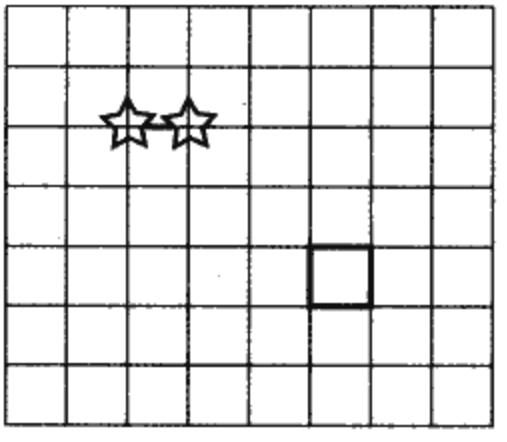


图 10.1　在格子上，夸克位于格线相交的地方，即一个格点上。图中用两个星形表示两个夸克。胶子位于格点之间的连线上。在图中，一个胶子连接着两个夸克，此外，还有一个纯胶子态组成的胶子圈

结果会出现一个非常复杂的计算问题。典型的格点数目由计算机运算能力的大小决定，例如时间和三维空间中都有 48 个点，总共就有 $48^4=5308416$ 个格点。由于把问题说清楚需要 32 个变量，计算总量相当于计算有 2 万多个面的盒子中的不规则区域的体积。可以通过告诉计算机采用“掷一下骰子”的蒙特卡洛方法将问题简化，使问题容易处理一些。然后对有限数目的夸克、反夸克和胶子组态平均，产生一个近似的求和值。通过“优化骰子投掷”的方法选择最重要的组态，计算这些组态的贡献。这些方法使问题的解决更加现实，但也付出了统计误差的代价。

格点 QCD 利用蒙特卡洛取样方法，靠蛮力完成对 QCD 路径积分的近似计算。微扰理论与它不同，需要通过越来越复杂的费

曼图一项一项地计算路径积分。

在实际操作中，要计算一个粒子的质量，比如说质子质量，格点模拟需要输入格点的数目，例如三个空间维度上每个维度 32 个点，时间维度一般会多一些，或许有 56 个点，还要输入夸克质量和相互作用强度。然后，要引入一个初始组态，可能是随机赋值的“热起点”，也可能是具有统一值的“冷起点”。然后将蒙特卡洛方法应用到该初始组态上，产生几百个“独立”组态。例如，很可能是产生两个上夸克和一个下夸克（即一个质子）沿着时间格子轴运动，这些都能通过对每种夸克、反夸克和胶子场组态的贡献取平均值计算出来。因为像强子这样的粒子在格子中传播的距离，会随着时间和距离的增加而衰退，较重的粒子运动的距离较短，可以利用“格点质子”在运动过程中衰变的可能性得到它的质量值。

得到的结果看起来十分令人信服。图 10.2 显示了一张用格点方式计算的粒子质量表格，质量较轻的强子的质量测量值用短横线表示。实际上，每种粒子在图中显示了两套不同的格点估计值。第一套用白方框表示，来自于唐纳德·温格特恩（Donald Weingarten）和他的同事在新泽西州约克镇高地的 IBM 托马斯·J·华生研究中心得到的结果。另一套数据（在图中用黑点表示）得到的时间要晚一些，来自于日本CP-PACS 合作小组。两组数据都令 K 介子的质量等于测量值，然后得到一个奇异夸克的质量，再用它计算余下的粒子质量。在近似理想的世界中，计算机得到的两套值完全一致；在完全理想的世界里，这两套值都与测量值吻合得特别好。实际上，格点 QCD 的结果同实际结果之间的差别不到几个百分点，因此，或许真实世界同理想世界之间的差别并不是很大。那么格点 QCD 理论已经证明了 QCD 理论的正确性么？它能够计算出我们想要的所有结果吗？当然，许多人对此十分激动。但是我们还是需要小心提防一些附加规则。

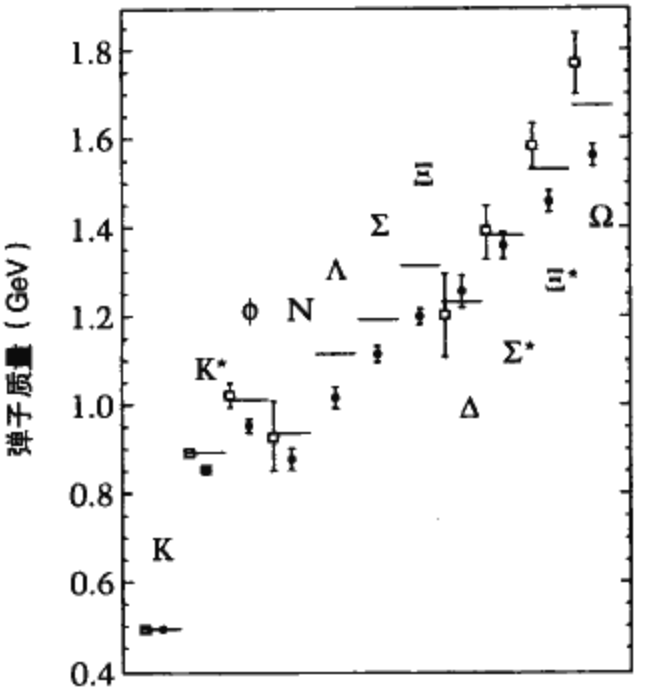


图 10.2　淬火近似下[1]得到的轻强子谱。短横线表示实验值，白方框是唐纳德·温格特恩得到的 GF11 结果，黑点是 CP－PACS 结果。图片来源：CP－PACS

格点生存的考验

“没有任何东西会成为问题，有的只是挑战和机遇。”这是一条由一位销售商提出的公司格言，他来自于一家世界上知名的大型计算机公司，该公司在格点 QCD 世界里也占有一席之地。这是一个充满机遇的世界，随便说两个例子，例如对夸克禁闭的验证，或者凭空计算出胶球的质量。各方面都不乏挑战。

其中一些挑战是格点方法固有的。在格点研究中产生无法消除的误差有两个来源，一个是在格点之间的间距大小有限，另一

① 译者注：quenched approximation，不考虑费曼图中费米子圈图的影响。

个（程度低一些）是格点时空体积大小也有限。计算机的存储能力和速度对这两者都有影响，尤其是格点间距应当越小越好。

格点间距是很关键的因素，因为任何从格点计算得到的有物理意义的结果——比如估计出的粒子质量——都必须与格点间距无关。否则，估计值就仅仅是格点的属性，而不是真实世界的性质。但只有当被模拟的物理问题，在格点间隔或者更小的尺度上没有重要尺度特征时，才能够满足这一要求。因此，如果想要得到精确而又有意义的答案，格点间距大小是关键因素。格点理论的研究者们一般都将格点间隔设置为大约 1/10 飞米。

由于可用计算机的运算能力大小决定了格点的总数量，因此格点间距同总模拟体积有关。这是因为，用格点表示盒子的总大小，就简单地等于格点数同格点间距的乘积。但盒子越大越好，一个直径大约为 2.5 飞米的盒子，差不多就能够确保将像质子这么大的粒子整个装进去。如果用一个比较小的盒子，盒子大小本身就会将问题扭曲。但如果盒子较大，格点之间的间距也会较大，这与保持格点间距最小的要求背道而驰。格点 QCD 充满了妥协，格点领域对越来越大及越来越快的计算机的要求永远也没有尽头。

除了格点大小的影响之外，利用格点也会导致一些令人头痛的问题。固体物理学家们了解用格点计算的知识。他们可以处理大自然为他们提供的格点——晶格。在晶格中，原子规则地排列在三维阵列中，一块晶体中包含上亿个原子。晶体物质的电性质和力学性质深受原子排列方式的影响。

由于晶格排列对原子的整体性质有影响，因此并不奇怪，令夸克和胶子呆在格点上会引入新的现象。实际上，它会造成一系列问题。仅仅因为格点的存在，就会引入额外的夸克——也就是那些“真实”夸克的复制品——这种情况偶尔会出现。正如威尔逊自己考虑到的一样，这些都是问题所在。为了解决该问题，威

尔逊在使用格点时将一个额外因子——现在被称为威尔逊项——引入了 QCD 理论的主方程中。该额外项能够将附加的多余夸克方便地除去。通过用这种方法改动主方程，威尔逊看到了解决格点方法固有问题的一线希望：当 QCD 理论的主方程为格点方法改写之后，会增加一整套新的项；但这些项反映的并不是真正的物理，因为将格点转换为真实的连续世界时，它们就会消失。威尔逊项就是这样一种附加物。

威尔逊项解决了一个问题——多余夸克的问题，但也付出了代价：它又引入了另一个问题。新问题的根源就是无质量夸克的手性对称性，我们在第九章中已经遇到过这一问题，左旋的夸克只能产生左旋夸克，右旋夸克也只能产生右旋夸克。即使对于没有质量的夸克，威尔逊项也会破坏这种对称性，并导致许多令人头痛的细节问题。特别是，夸克质量在格点计算中开始变得像变魔术般不可思议，只能通过在模拟过程中凭经验仔细地手动调整参数才能将其除去，这是人们不愿意看到的。

威尔逊沦为了一条基本规则的牺牲品。该规则是，不可能建立一套单夸克格点 QCD 理论，在不破坏手性对称性的前提下，保持所有重要的颜色对称性。有一些模拟程序使用了威尔逊的方法，但还有其他一些流行的选择。在伊利诺伊大学（位于乌尔班纳）的约翰·科古特（John Kogut）和斯坦福大学的伦纳德·萨斯金德（Leonard Susskind）设计了一套方法，用这种方法会让 QCD 在格点展开中满足手性对称，但其代价是每一个真实夸克味都要对应 4 种夸克。从本质上讲，就是一个单独的夸克要散布在 4 个邻近的格点之上。这种方法不存在必须人工剔除夸克质量的问题，并且利用他们称为“错列费米子”（staggered fermion）的方法可以很自然地得到几乎无质量的夸克。它虽然在技术上有一些不利因素，最明显的是 4 种错列格点夸克之间的关系，即上夸克、下夸克、奇异夸克和粲夸克之间的关系。这看起来有点混

乱，但现在主要的大规模 QCD 模拟都是用错列费米子方法实现的。发展和改进这种将 QCD 主方程散布在格点上的方法，在格点领域是一个热门课题。

在格点方法中，各点之间的间距以及容纳夸克和胶子的盒子体积有限，是带来系统误差的两个主要方面。“系统误差”的意思是，给定了格点间距和盒子大小之后，计算机无论计算多长时间都无法给出更好的结果。另一个误差来源是由蒙特卡洛方法带来的统计误差（计算机运行更长时间可以减小该误差）。这三种误差都会在所有的格点计算中出现。雪上加霜的是，还存在第四种误差，它本质上存在于现有的各种格点计算中，实际上它是被格点理论物理学家们故意引入的。它被称为淬火近似。

让一次格点计算囊括所有可能的夸克贡献，会耗费计算机大量的时间。这部分是因为，在计算过程中会产生大量夸克-反夸克对，但主要还是因为夸克和反夸克必须遵守的泡利不相容原理而带来的大计算量。1981 年，温格特恩和罗马大学的吉尔吉奥·帕利希（Giorgio Parisi，图 10.3）巧妙地简化了问题，他们简单地用数字 1 替代了烦人的复杂部分。这就是淬火近似。尽管只做了很小的调整，似乎是一种权宜之计，但这种办法却简单而有效。从物理上讲，淬火近似——有时也被称为价近似——实际上是忽略了所有从胶子和真空中产生的虚夸克-反夸克对。换句话说，海夸克都被忽略掉了——就好像它们的质量为无穷大——只有价夸克能体现出来。

好的方面是，研究格点的人们利用淬火近似方法可以在合理的时间内得到答案。实际上所有早期的格点计算都用了淬火近似，在图 10.2 中所示的那些数据也用了该方法。不好的方面在于这些结果的准确度不太高。例如日本 CP-PACS 合作小组在 1998 年得到的结果（图 10.2）在 10%的误差范围内同已知的粒子质量相符。格点理论的研究者们知道，这些误差中大部分是由

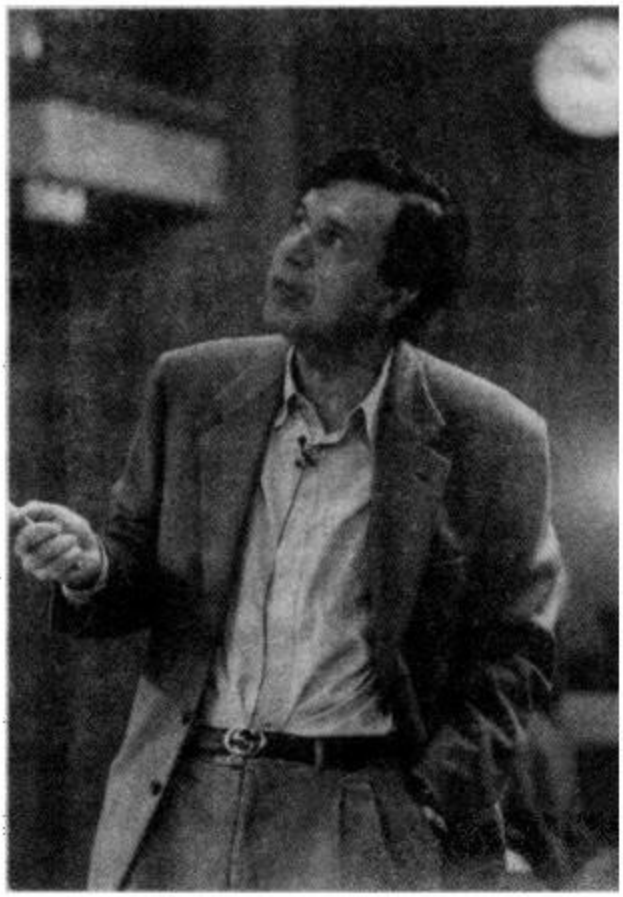


图 10.3　吉尔吉奥·帕利希 2001 年在巴黎的一次会议上演讲。帕利希的兴趣不限于 QCD 理论，还包括很多其他领域，如聚合、玻璃体、其他无序系统以及非平衡物理。图片来源：约翰·埃利奥泡洛斯

于淬火近似引起的，他们现在对这种结果无法接受，或者至少认为太粗略了。这个问题最直接的解决方法就是用动力学夸克进行所谓的“完全 QCD”计算，换言之就是允许所有的夸克都可以成对产生和湮灭。问题在于，同淬火近似相比，完全 QCD 计算需要几百倍以上的计算能力。大多数研究中心的格点研究者们，现在都已经从淬火近似转移到完全 QCD 计算上来。他们已经得到了轻夸克质量的改进结果，误差精度达到了百分之几，并且还有可能大幅改进。他们所需要的只是巧妙的算法和非常非常快的计

算机。

速度的需求

“在研究中，你需要具备三样东西：金钱、金钱还有金钱。”——这是一句谚语。当然巧妙的思想也很有帮助，但在格点 QCD 中，只有思想是远远不够的。你必须有资金建造足够强大的计算机来完成和测试那些巧妙的算法，并且在合理的时间之内产生有意义的结果。

一台家用的个人电脑或者是研究者的台式工作站都无法满足格点 QCD 的需要。格点研究者们渴望拥有最大并且最快的计算机，他们用的机器通常都是定制的，供他们的工作专用，通常可以将负荷分配给统一调度的许多独立处理器。它们都是并行的超级计算机。

首批这种 QCD 计算机中的一台就是由温格特恩和他的同事在 IBM 建造的。他们把这台机器叫做 GF11。它拥有 566 个可用来并行处理同一个问题的处理器，本质上就是 566 台一同工作的计算机，每秒钟能够进行 110 亿次运算，这就是它的名字 GF11 的由来。计算速度的单位是每秒的浮点操作数，它是所有高性能计算科学讨论中都会涉及的关键参数，粗略地讲，它就是机器在一秒钟能够计算两个大数相乘的次数。由于计算机的运算速度特别快，因此更加常用的单位是兆次（megaflop），即 Mflop，它表示“百万次浮点操作每秒”。温格特恩建造的那台机器运算速度达到了上 10 亿次，该单位叫做吉次（gigaflop），或简写为 Gflop。因此，GF11 是一台运算速度能达到 11Gflop 的机器，见图 10.4。

从 1983 年开始，人们耗费 8 年时间设计和建造了 GF11。这台机器调动 80%的运算资源进行 2 年的计算之后，得到了胶球的质量和寿命。根据 GF11 得到的结果，无自旋的胶球质量大约为

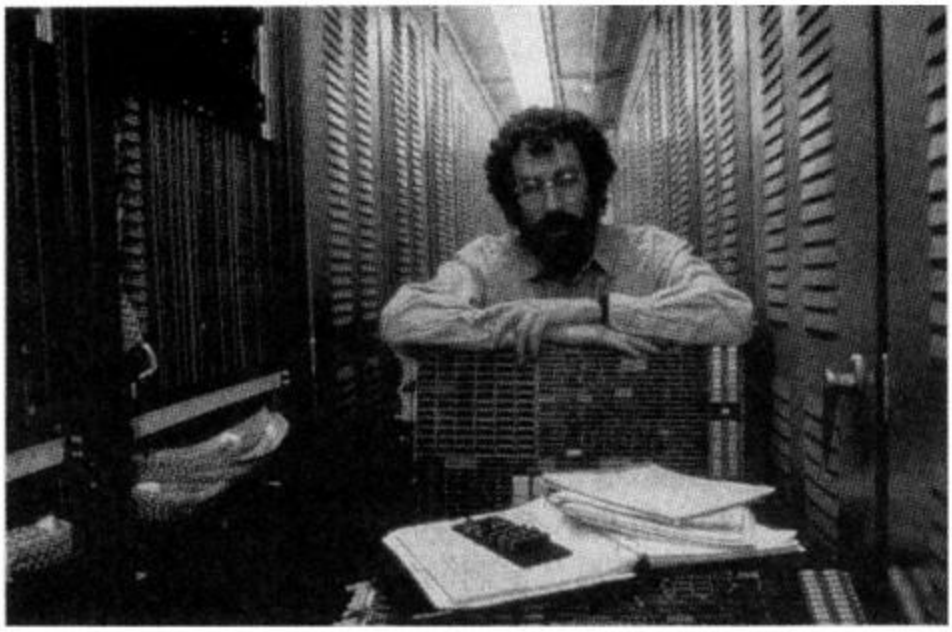


图 10.4　人与机器。唐纳德·温格特恩同 GF11 在 1993 年一起拍摄了这张照片。温格特恩身后的两排“档案柜”中，每个柜子里都安放着半台 GF11 的处理器。左边，处理器架的一个盖子已经打开了。图片来源：IBM

1740 兆电子伏，该质量与胶球候选者 $f_0(1710)$ 的质量十分接近。GF11 小组在 1994 年得到的轻夸克质量谱——见图 10.2——已经变成了其他小组检验自己结果的参照标准。

这些小组之一就是日本筑波大学的 CP-PACS 小组（Computational Physics by Parallel Array Computer System，并行阵列计算系统计算物理小组），他们所得的数据也列在图 10.2 中。CP-PACS 计算机还进行天体物理和凝聚态物理方面的计算，尽管它没有完全服务于格点 QCD，但在 20 世纪 90 时代末，它仍是进行 QCD 计算的最强大计算机。它拥有两千多个独立的处理节点，这些节点基本上都是速度更快的 Hewlett Packard 商用计算机，这些计算机在三维空间中按 17×16×8 的立体阵列排列，不同节点之间采用了与众不同的连接方式。CP-PACS 计算机的理论处理

能力是614Gflop，它在运行测试程序时的速度达到368.2 Gflop，这暗示了这些大型计算机的一个问题——它们处理实际问题的速度有时会远远低于所标称的速度，或者说是理论计算能力。另一个小问题是它们用起来不方便，一些非常大型的计算机使用起来很麻烦，因为它们要用特殊的计算机语言。但CP-CAPS计算机能够有效地运行用标准高级语言编写的程序，例如Fortran、C和C++，只有最里层的循环使用特殊的“汇编”代码，以实现最高的速度和效率。这台机器造价2200万美元，它的外形看起来像一座装满单调乏味而又昂贵的文件柜的房间。CP-CAPS计算机是它这一代计算机的典型产品，是由物理学家和计算机科学家组成的大型合作团队共同设计的专用产品。现在的物理学家比以前任何时代的科学家，都要更精通于精密的机器制造。

这些大型QCD计算机散布在全球各地。位列QCD研究榜首的是具有12288个节点，运算速度达到180Gflop的QCDSP机器，它位于布鲁克海文国家实验室的RIKEN布鲁克海文研究中心。在纽约的哥伦比亚大学还有一台具有8192个节点、运算速度达到120Gflop的QCDSP机器。APEmille机器阵列为廉价大型计算机设立了一个新标准，它具有250Gflop的运算速度。在INFN，即意大利罗马的国家核物理研究所，有两台128Gflop的机器，DESY也有一台250Gflop的机器。

QCDSP和APEmille都是自行建造的，它们都基于大量相同的专用处理节点，连接成奇怪的阵列，使机器具有格点计算的能力。另一些重要的机器使用了大量商业计算机。其他一些小组走了另一条路——工作站农场。这种想法十分简单，就是将大量工作站或者PC机捆绑到一起，用普通的电脑网络将它们连接起来，将计算过程分布在各个计算机上进行。农场还无法同专业机器的运算速度相比，但它们更加便宜和灵活，并且很容易利用计算机技术的快速发展加以改进。

随着时间的推移，我们迎来了新千年的曙光，只有3台计算机突破了太次（Tflop）的心理壁垒，全世界最强大的计算机就是英特尔公司的ASCI Red，它位于新墨西哥州阿尔伯克基市的圣地亚国家实验室。这台机器的运算速度达到了2.379Tflop，本来用于核武器的模拟测试，但它仍不能满足格点QCD领域对计算机能力的需求。2004年，世界上运算速度最快的QCD计算机在爱丁堡问世，它就是UKQCD合作小组的QCDOC，它的名字代表“芯片上的QCD”（QCD On a Chip）。这是一种为QCD计算特意定制的芯片，这台机器和它在布鲁克海文的姐妹机的运算速度达到了5Tflop。这并不是终点——更强大的机器也即将出现，谁将是第一台达到拍次（Petaflop，即千万亿次）运算速度的计算机呢？

格点的成果

利用轰鸣的计算机计算出一些粒子的质量，然后将这些质量列在数据表中供所有人查阅，看上去并不十分惊险刺激。但是算出轻强子质量谱，在粒子物理学家心中永远是一块心病。原因很简单：能够从纯理论出发，计算出质子和其他一些普通粒子的质量，是粒子物理的伟大目标之一。物理学家们实现该目标的那一天，就是他们真正理解和掌握世界在最小层次上运行规律的一天。QCD理论在低能情况下会遇到特殊的挑战，从头开始计算一个复合粒子质量的过程能够体现出这种艰难和理论的苍白。还有一个更实际的原因——人们都知道极其精确的质量值，这显然是检验格点QCD理论的好办法。自从格点QCD理论诞生以来，强子质量谱就成为新思想和新算法试车的地方。

格点研究者们还估算了许多其他的量。例如，计算出了一个强耦合参数的值以及4种最轻的夸克的质量，还有底夸克质量非

常精确的估计值。格点计算还能够清楚地展示在第七章中遇到过的“$\Delta I=1/2$ 定律”，其中强相互作用限制了弱衰变。一种别在领口的徽章就取自于对 D_s 介子的研究，这种介子由粲夸克和反奇异夸克组成（D_s^+），或者由反粲夸克和奇异夸克组成（D_s^-）。D_s 介子的衰变常数可以衡量它经过弱相互作用发生衰变的概率，格点计算预言的衰变常数值为 220±30 兆电子伏，尽管这个值已经打上了淬火近似的标签，但它与实验值 241±32 兆电子伏符合得很好。格点研究者们对计算 B 介子系统的性质以及能提高 B 介子物理精确程度的实验工具特别感兴趣。

或许最令人感兴趣的格点计算是那些胶球和混合介子的计算。根据淬火格点计算，应当存在一整套胶球质量谱，最一开始是自旋为 0、质量近似为 1.7 吉电子伏的胶球，接下来的胶球自旋为 2、质量大约是 2.4 吉电子伏，另外还有许多具有各种量子数、质量更重的粒子。但正如在第六章中讨论过的那样，找到胶球是一项很难的任务。寻找胶球的一种通常做法是寻找具有“奇特”量子数的胶球，换言之，对应一个物体的量子数超出了普通夸克模型的框架。根据格点 QCD 的预言，这种态中最轻的也超过 4 吉电子伏，这个质量即使以胶球的标准来看也太大，这或许也是它还没有被实验者们观测到的原因。混合介子——我们在第六章中也介绍过了——就是像 $\pi_1(1400)$ 粒子这样的物体，它是 $q\bar{q}g$ 的组合，这种组合方式已经超出了普通夸克模型。在第六章中提到的计算机模拟实际上就是格点 QCD 计算，具有与 $\pi_1(1400)$ 粒子相同量子数的最轻的奇特介子，质量应该在 1.8～2 吉电子伏，这个值太高了。使胶球和混合介子的格点结果同真实世界一致，仍然是一个吸引人的挑战。

计算机禁闭……

虽然一些课题——比如计算核子结构函数以及夸克和胶子碎

裂为喷注等过程，尚处在萌芽阶段，但格点 QCD 的潜力仍十分巨大。格点是能够透彻了解瞬子的唯一方法。但什么才是最终的奖励？夸克禁闭？

前些年，许多事情看起来都很有希望。在 20 世纪 80 年代，格点计算结果显示，夸克同反夸克之间的相互作用能量会随着两者间距的增加而必然正比增长。这样，就证明了夸克禁闭。可是，还不确定。困难看上去是细节性的，但因这一点极其重要，而让我们无法打开香槟举行庆祝。格点 QCD 得出的夸克禁闭结论，来自于一个相对粗糙的格点。格点夸克之间的间隔很大，因此它们之间的相对作用强度也较强。要将其变为真实情况，格点间距必须缩小到零，问题就出现了。随着格点间距的减小，夸克-反夸克之间的间距也随之减小，这就表示夸克之间的相互作用强度也变得越来越小，渐近自由起作用了。但是在缩小间距的过程中却丢失了夸克禁闭性质成立的证据——因为相互作用强度减小，这种性质不能再理所当然地成立。尽管计算机模拟表示，随着格点间距减小，夸克禁闭确实仍然存在，但没有人能够证明这一点。有可能这些都是理论上的结果，我们有必要回想一下，QED 理论在粗糙的格点中也显示出禁闭的性质。如果 QED 理论在格点间隔为零的时候，夸克禁闭仍然存在，将会产生一场灾难。这说明电子同夸克一样也会禁闭，因此，就不会存在我们熟知的电视这类东西了。但在 QED 的情形中，研究者们能够说明，从粗糙格点的禁闭 QED 到我们所在的非禁闭平滑世界间存在一次跳变，或称一次相变。我们所面临的挑战，就是要确定无疑地证明 QCD 理论并不存在这种跳变。

但格点模拟还可以用其他方法帮助我们解决禁闭难题。一些研究者声称，格点结果支持在第七章中已经介绍过的双超导禁闭模型。这种模型是将介子中连接夸克和反夸克的色-电力，挤压到环绕着单极子海的管子中。因此单极子凝聚态会将色-电力线

排斥出真空，因而形成了禁闭力管。这样，禁闭实际上就是从普通真空转化为超导真空。根据这幅图像，夸克周围环绕着一个小的真空区域，或者称真空气泡——这一思想同重子口袋模型呼应。当介子中的夸克和反夸克接近，它们各自的普通真空气泡会重叠。它们之间的作用力由于受到反屏蔽作用的影响，会处于最小值。如果将夸克和反夸克拉开，普通真空泡会变成由超导真空环绕的哑铃形普通真空，哑铃的颈就是连接夸克对的色力管。

……以及退禁闭

还可以用另一种方法考虑整个禁闭问题。只要涉及海水、陨石、铌球等日常世界中的事物，夸克（和胶子）都会禁闭。但它们一定总是禁闭吗？夸克和胶子能否在一些特殊环境下成为自由粒子？如果这样，有没有可能引起一次跳变，将禁闭世界变为一个非禁闭的世界呢？

粒子物理学家们相信自由夸克和反夸克可能存在，并且这种跳变可以实现，但必须在十分特殊的环境下。当像质子和中子这样的普通粒子被强行挤压到一起时，它们会失去各自的特征，就像挤在轧酒机中的许多葡萄一样。随着复合粒子相互重叠，每个粒子的内部组分就会暴露给其他粒子的组分。由夸克组成的质子呈颜色中性，但现在它们的角色也可以用来自于其他粒子的夸克来完成。在这种环境下，单个强子的概念变得多余，它可以被一种自由夸克和胶子气体，或称“等离子体”的概念取代。

尽管渐近自由对这种夸克-胶子等离子体中的夸克解放起一定作用，但由于胶子屏蔽了夸克之间的色力，理论物理学家们还是相信有一种新机制在起作用。我们还需要从固体物理世界借用模型，在原子物质中，带正电的离子产生的电场被它周围的电子修改。由于电子带负电荷，这些环绕电子会对带正电的离子产生

部分屏蔽。当距离比正离子之间的距离短时，屏蔽强度足以压制正离子的影响，一些电子就会变成在材料中四处游荡的自由电子，形成一种能够导电的“电子海”。但如果屏蔽是长程的，电子仍然会束缚在离子周围，材料就是一种绝缘体。但如果绝缘体中的屏蔽范围能够相对离子间距进行调整，离子就有可能放走一些束缚电子，从而使绝缘体转变为导电的金属。内维尔·莫特（Nevill Mott）在1968年首次提出了这种绝缘体-金属变换，这种转变现在已经很常见了，可以通过一系列效应产生，特别是通过加压。压力会减小离子之间的间距，增加电子的聚集程度。这说明屏蔽效果会更强。乍一看，好像不会发生任何变化，因为缩小离子之间的间距，束缚电子会更紧地束缚在离子周围。但是电子密度的增加有助于屏蔽效应的增强，增强的程度要比空间缩小程度强得多。屏蔽效应获胜了，因此挤压能够使绝缘体变为金属。例如在1.4兆大气压和2300℃下，氢会变成金属并能够导电。

绝缘体通过电荷屏蔽效应转变为导体这一现象，为我们提供了从强子经过颜色屏蔽转变为夸克-胶子等离子体的模型。随着像质子或中子这样的粒子被挤压到一起，每一个夸克都会发现，自己正同其他夸克摩肩接踵地挤来挤去，而并不只接触来自同一个粒子的夸克。但挤压会增加胶子的密度，类比于莫特转变中电子密度的升高，它会将任何一个夸克周围的那些夸克屏蔽掉。这样，那些开始只是某个复合强子一部分的夸克（或反夸克）就会发现自己从混战中摆脱出来，几乎不能感觉到各自的存在。同自由电子导致导体中产生电荷流动十分类似，这种夸克-胶子等离子体也能产生色流。非禁闭的情况实际上可能更加简单。伯利亚科夫和萨斯金德（Susskind）都认为，将夸克束缚在一起的弦在高温下会简单地熔化掉。

格点计算表明，从可辨认的粒子态（例如质子、中子等）到另一种新物质态（例如夸克-胶子等离子体）的过程会发生跳变，

或者称相变，这种跳变非常突然。由于这种相变是从一个夸克和胶子禁闭的区域转变为一个不禁闭的区域，因此它被称为退禁闭相变。从禁闭相跳变到另一个非禁闭相的证据，当然也是一种跳变的禁闭相存在的证据。根据格点研究者们得到的结果，退禁闭跳变应该发生在大约 10^{12}℃，这大约是太阳核心温度的 10^5 倍，其密度要达到普通物质密度的 10^{20} 倍，是普通原子核密度的 10 倍左右。所有能够使强子相（包括普通的核物质）跳变为夸克-胶子等离子体的温度和密度结合起来，将会在温度-密度图上描绘出一条“相边界”。在相对较小的密度下达到等离子态需要极高的温度，如果温度较低则需要极大的密度。

仅仅到达等离子态可能还不是故事的结尾。除了退禁闭跃迁之外，还有另一个层次的结构需要考虑。在第九章中我们提到过，铁磁体是自发对称性破缺的原形。大量原子磁体的排列形成了能量最低态，这种态缺乏原子磁体基本相互作用的转动对称性。至关重要的是，它发生在某种特定的温度之下——如果升高磁体的温度，方向一致性就会消失。这时，磁体就会从铁磁态跳变到更像普通金属的态。

设想一下对 QCD 真空进行相应的操作。如果向真空中注入足够能量，夸克凝聚态中夸克和反夸克之间的联系将会随着凝聚态“熔化”而遭到破坏。随着夸克凝聚态消失，所有手性对称性破缺的特征也会消失，手性对称性得以恢复。这样，除退禁闭相变之外，在高温下还会出现另一种相变，它能够引起理论物理学家们称为手性恢复的过程，左旋夸克和右旋夸克老死不相往来的手性体制再次出现。在质子和中子这类普通粒子的手性对称破缺相中，由于发生手性对称性破缺，夸克具有大约 300 兆电子伏的有效质量或组分质量。在手性对称性恢复温度之上，手性对称相中夸克的质量近似为 0：它们的质量仅仅是所谓的当前质量值，上夸克和下夸克的当前质量分别为 4 兆电子伏和 7 兆电子伏。理

论物理学家们预言，由于手性恢复的结果，一些由夸克-胶子等离子体产生的粒子，其质量、寿命以及衰变产物谱都会不同于现在的粒子表。但退禁闭同手性恢复转变是否有联系，甚至它们是否能够在完全相同的条件下发生，仍然不太清楚。

最近，对解释手性对称性破缺十分重要的夸克凝聚思想，在另一个问题中变得越来越重要。在上面提到的通常的真空图像中，左旋夸克与相同味的右旋反夸克凝聚在一起，反之亦然。但是在1999年，弗兰克·维尔茨克（Frank Wilczek）、马克·阿尔弗克（Mark Alford）、克里希纳·拉亚戈帕尔（Krishnas Rajagopal）和托马斯·谢费尔（Thomas Schäfer）提出了一个完全不同的大胆设想。他们认为，在非常致密的核物质中，夸克配对的方式更像普通超导体中的电子对，结果就形成了以“夸克库珀对”为基础的“颜色超导体”。

物质的核心是一种新的、由不同味和色夸克组成的夸克凝聚态。现在这是一个有争议的问题：在高能物理中，通常完美的颜色对称性要高于不那么完美的味道对称性，但这里却不同。新的凝聚态将两者混到一起，并被称为“色-味关联”（color-flavor locking）。

推理过程如下：两个夸克在互相接近的时候，总色荷减少，产生吸引效应，这种效应同固态超导体中晶格在电子间传递的那种吸引效应类似。同在大街上走过的朋友一样，反向行进的夸克会发现它们以夸克库珀对的形式四处游荡。这种协同行为——或者说凝聚到颜色超导态的效应，在整体能量方面有优势，这非常像固体超导态中真实的库珀对。

这种真空会带来大量深远的影响。孤立的色和味立即不再是好的对称性，只保留下某些复合对称性——那些锁在一起的对称性。胶子通过迈斯纳效应得到了质量，这特别类似于光子通过固态超导体时获得有效质量的过程。粒子的电荷也发生了改变，有

一些胶子甚至获得了电荷。实际上，由于电荷都取整数值，并且基于色的长程力由于胶子具有质量而消失，理论就会产生关键的禁闭效应。另一个具有很高价值的结果是手性对称性破缺，它也被照顾到了，右旋味对称性与色关联，因此也与左旋味对称性关联，左旋和右旋变换不再独立。一切都正常了，按照这套理论，在高密度极限下的性质，同真实粒子组成的低密度世界相符。或许在日常世界里，质子和中子中的夸克运行情况是这种超密态的某种极限。到目前为止，色-味关联还缺乏实验和模拟的支持，但它的引入在理论物理学家中掀起了一阵热潮。

创造一种新的物质态

粒子物理学家们相信，在大爆炸后最初几微秒的时间内，大自然能够提供夸克-胶子等离子体所需的环境。一种重现该环境的方法，就是将几束夸克和胶子撞到一起，产生一个微型火球，在地球上形成一次“小爆炸”。该过程能够产生一串夸克-胶子等离子团，这些等离子团会发生膨胀和冷却，同时析出 π 子和其他一些粒子。实验物理学家们认为，他们能够造出这样的等离子体，并探测到析出的粒子，从而证实他们的推断。

纽约布鲁克海文国家实验室的物理学家们在格点 QCD 计算的引导和鼓舞下，试图在很小的范围内再现宇宙诞生最初几分钟的情况。1999 年，这里的实验者们在相对重离子对撞机（Relative Heavy Ion Collider，简称 RHIC）上，以运动非常快的重离子形式将夸克束或胶子束撞到一起，希望能够产生真正的夸克-胶子等离子体。

RHIC 是一种新型的对撞机，它能够将高能粒子束（而不是像电子或正电子这样的单个粒子）集中在一起。它利用已经取消的伊萨贝尔质子-正电子对撞计划建造的隧道，并且用现有的

AGS 加速器做注入器。在周长 3.8 千米的环上，有大约 1600 块超导磁铁对离子束进行弯折和聚焦，可加速的离子可以达到极重的金离子。原子在经过很薄的金属箔片时，被夺走了电子，形成带有高电荷的离子，它们同迎面而来的离子束带电符号相同。这种反应的第一个特点是，离子束内的离子很容易互相排斥，使得离子束变得很宽，因此 RHIC 的一个特殊性质就是它具有较宽的离子束管道，直径为 8 厘米。第二个特点是两离子束具有相同电荷，因此会有两束完全反向的离子束，在它们各自的粒子束管道中在磁铁的控制下旋转。每个核子能量高达 200 吉电子伏的粒子束发生碰撞，有 4 个探测器探测这些碰撞：PHENIX，BRAHMS，PHOBOS 和 STAR。在寻找夸克-胶子等离子体时，每检测到一次离子-离子碰撞，这些探测器都要面对多达 1 万个粒子产生的洪流，其中大多数是 π 子。

实验物理学家们寻找等离子体的一种方法，就是寻找在重离子碰撞中所产生的 J/ψ 粒子数量下降现象。德国比勒费尔德大学的赫尔默特·萨斯（Helmut Satz）和麻省理工学院的天津雄松井（Tetsuo Matsui）早在 1986 年就提出了这种想法。这种想法认为，粲夸克和反粲夸克在火球中以夸克-反夸克对的形式产生，并且相互飞离。随着等离子体的冷却，粲夸克和反粲夸克并不会结合为 J/ψ 粒子，它们会在运动过程中同较轻的夸克配对，成为可观测的粒子。因此，如果这个机制是正确的，在实验条件能够产生夸克-胶子等离子体的那一刻，J/ψ 粒子的产生率将会突然下降。轻夸克和反夸克将会以同样方式失去它们的同伴，并且很容易地找到替代者。粲夸克的特殊之处在于它的质量很大，因而数量相对稀少。

早期的格点计算表明，CERN 的 SPS 应该能够产生合适的高温和高密度，从而生成等离子体。实验者们从 1986 年开始在 CERN 进行这项实验。他们一开始利用的是硫离子，观察在理论

物理学家们认为不可能触发等离子体的条件下，会有多少 J/ψ 介子产生。在 1994 年，SPS 升级为使用铅离子，理论物理学家们认为，这些较重的弹射物能够在碰撞中产生更大的能量密度，足以让人们看到等离子体。由 CERN 的 NA50 小组和其他一些小组收集的数据，确实显示出 J/ψ 粒子产量下降的现象。实际上到 1997 年，这种下降已经足以说明夸克-胶子等离子体很可能已被证实了。这些数据的关键特征在于，不仅存在下降，这种下降还特别急剧。人们认为这是由于物质转变为一个新物质态而形成的，这种急剧下降不利于更加普通的、对 J/ψ 粒子产量下降的“吸收”解释，吸收解释预言的 J/ψ 粒子信号下降更加平缓。CERN 的实验者们还见证了另一些暗示存在等离子体的信息，其中包括奇异粒子产量的增加。特别是 Ω 粒子的产量可以达到“普通”产量的 15 倍，在质子-质子碰撞过程中就是这样。即使是这样，粒子物理学家们还是不敢轻易断言已经发现了夸克-胶子等离子体。专门应夸克-胶子等离子体要求而建造的 RHIC，将会决定这种等离子体是否真实存在。

RHIC 能够将足够多的能量挤进足够小的体积内，从而进入等离子的范围，产生足够大和寿命足够长的等离子态，最终完全揭示它的规律。等离子体的一个信号就是它的大小，其直径能达到 40 飞米，比产生它的离子还大几飞米。物理学家们认为，这种膨胀由最初的热导致，并且同等离子凝结为无序度降低的可观测粒子的过程相反，这种膨胀见证了热力学规律要求的混乱程度整体升高现象。实验物理学家们计划用从该体积内发射大量 π 子的相对时间来获得等离子体的大小，实际上他们用 π 子重复了天文学家用光测定恒星大小的方法。这些 π 子还能够提供等离子体的另一个信号，即随着实验能量的提高，超过等离子产生的极限，π 子的数量会发生激增，见图 10.5。

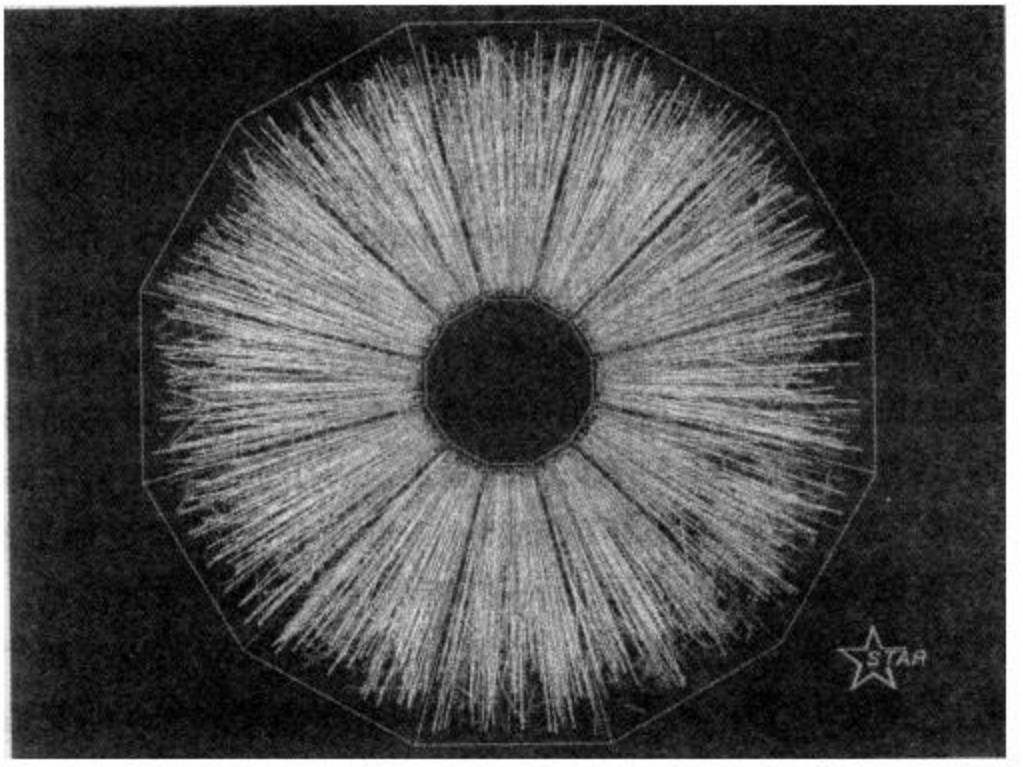


图 10.5　分析一下这幅图。在 RHIC 上进行的金离子束碰撞，打破了一次亚原子碰撞产生粒子数的纪录，这是由 STAR 探测器记录下来的。RHIC 的科学家们正在搜寻夸克-胶子等离子体存在的信号。图片来源：布鲁克海文国家实验室，STAR 合作小组

太空，最后的前沿

一颗恒星的质量如果为我们太阳质量的 2～3 倍，它的核燃料最终耗尽之后会开始冷却。这座核地狱曾经产生的、足以支持恒星膨胀的内部压强已经消失了，但引力仍然存在。在引力的作用下，组成恒星的粒子相互挤压，越来越紧密，直到质量巨大的恒星坍缩为一个叫做中子星的椭球体，它的直径只有几千米。中子星的密度十分巨大。顾名思义，中子星主要是由中子构成的，

质量更大的恒星将会形成一个黑洞。

旋转的中子星可以通过它发出的强烈电磁辐射被观测到，这些电磁辐射会扫过整个宇宙，就好像从灯塔中发出的灯光扫过海洋一样。在地球上，这种电磁波束呈现为有规则的脉冲，旋转的中子星就是一颗脉冲星。

中子星的赤道由于自转而膨胀，随着年龄的增加，它的能量会流失到太空中，因此它会逐渐慢下来，由于自转的离心效应变小，会变得瘦一些。对于一些旋转的中子星，随着自转变慢，由于密度增加而引起的额外挤压会将内部的核物质转化成一种冷夸克-胶子等离子体，称为夸克物质。这就是维尔茨克和他的同事在探索高核密度 QCD 时想出来的东西。正如布鲁克海文国家实验室的诺曼·格兰德宁（Norman Glendenning）和他的同事提出的，产生“夸克星”的一种可能信号就是脉冲星脉冲信号的忽然中止，否则它的频率会慢慢下降。

根据估计，每 100 个脉冲星中会有 1 个经历这一过程，天体物理学家们现在已经参与了搜寻退禁闭物质的比赛。甚至宇宙学家们也对夸克-胶子等离子体很感兴趣，因为如果它确实在我们宇宙的早期阶段出现过，它留下的烙印将会通过引力波渗透在空间中。在本书的开头，由于同 QCD 的故事无关而丢掉的引力，可能通过恒星的引力坍缩或者整个宇宙颤抖的时空涟漪，回到我们的故事当中。在统一的物理学中，不存在漏网之鱼。

附录 1 QCD 年表

1864 麦克斯韦统一电和磁，并提出光是一种电磁波的假设

1896 贝克勒尔发现放射性现象，它是弱相互作用力（弱力）的结果

1897 汤姆逊发现电子

1900 普朗克发现宇宙的量子本质

1905 爱因斯坦用光量子解释了光电效应

1905 爱因斯坦提出狭义相对论，这是一种可应用到速度接近光速物体的理论体系

1909 盖革和马斯顿观测到 α 粒子在金属箔片上的大角度散射

1911 卢瑟福将盖革和马斯顿观测到的结果解释为原子具有原子核，并且从带电粒子同点状原子核散射，推导出他的著名公式

1911 密立根测定电子电荷

1912 赫兹发现宇宙射线

1912 威尔逊发明云室

1913 玻尔提出原子模型：原子由原子核同围绕原子核的轨道电子组成。他解释了氢光谱，并提出放射性是原子核的一种性质

1918 威尔为了尝试将引力同电磁力统一起来，提出“规范不变性”这个术语，并发现了它同电荷守恒之间的联系

1918 诺特定理（Neother's theorem）将对称性和守恒量联系起来

1919 卢瑟福将氮转化成氧，并且首次揭示质子存在的证据

1920 卢瑟福推断质子是原子核的一部分，并且猜测存在中子

1921 查德威克和比勒尔在 α 粒子同氢原子核的散射实验中，首次观测到强相互作用力（强力）的证据

1923 德布罗意提出电子的波粒二象性

1925 海森堡提出矩阵量子力学

1925 泡利提出不相容原理

1925 古德斯密特和乌伦贝克提出电子具有自旋，以解释原子光谱中观测到的双线

现象

1926 薛定谔写出了量子力学的波动方程

1926 魏格纳将群论引入量子力学

1927 海森堡提出不确定性原理

1927 戴维森、革末和汤姆逊利用晶体衍射发现电子的波动行为

1927 安索尔德利用薛定谔量子力学计算氢的能级

1927 福克将规范变换的现代形式引入了 QED 理论

1927 狄拉克提出量子电动力学和量子场论

1928 伽莫夫、伽尼和康顿用量子隧道效应解释 α 粒子衰变

1928 相对论性质的量子理论——狄拉克方程被提出

1928 维德罗意建立了世界上第一台共振加速器，它是回旋加速器和同步加速器的前身

1928 威尔概括出规范不变性同量子理论中的电荷守恒之间的联系

1929 威尔采用福克的局部相位修正，但仍保留了“规范不变性”的名字

1929 格思森进行首次质子-质子散射实验

1930 泡利提出存在中微子

1930 莫特在卢瑟福的基础上提出了一套公式，用来描述具有自旋的粒子散射。例如，在电子同点状核子的散射过程中应该计入电子自旋

1931 狄拉克在量子理论的情形中讨论了磁单极子

1931 劳伦斯和利文斯通发明回旋加速器

1932 贝特和费米提出相互作用力的粒子交换机制

1932 查德威克发现中子

1932 尤里、布里克怀迪和墨菲发现氘

1932 考克饶夫和沃尔顿击碎原子

1932 海森堡提出，原子核由中子和质子组成，并且认为它们是一个粒子的两个荷态

1933 卡尔·安德森发现正电子

1933 狄拉克发表关于波函数变换的论文，为费曼的路径积分方法提供了关键基础

1933 布莱克特观测到电子-正电子偶的产生过程

1933 斯特恩测量质子磁矩

1933 费米提出 β 衰变的理论，同时首次提出自旋为 1/2 场的二次量子化

1934 斯特恩测量中子和氘核的磁矩

1934 费米开始进行中子散射的大规模实验

1935 怀特公布首次“高能”加速器质子-质子散射实验，并说明质子行为不同于点电荷

1936 塔弗等人利用自动粒子探测，极大地改进了质子-质子散射实验

1936 布赖特、康顿和布莱森特提出质子-质子散射和质子-中子散射过程中的相互作用力相同

1936 康顿和凯森、魏格纳（1937）分别提出质子和中子同位旋的现代形式

1936 布赖特和魏格纳发表共振公式（Resonance formula）

1937 卡尔·安德森和奈德梅耶发现 μ 子
1937 在斯坦福发明了速调管，它是一种强大的微波的发生器
1938 哈恩、斯特拉斯曼、迈特纳和弗里希发现原子核裂变
1940 泡利证明粒子旋转同统计学之间的关系
1942 费米建立第一座核反应堆
1944 路易斯和迈克尔公布存在带正电的 K 介子——第一种奇异粒子——的首次证据
1944 维克斯勒和麦克米伦（1945）提出加速器的相位稳定性
1946 第一台同步回旋加速器（能将质子加速到 350 兆电子伏）在伯克利开始运行
1947 鲍威尔和帕金斯发现带电 π 介子
1947 兰姆和卢瑟福测量兰姆位移
1947 罗彻斯特和巴特勒发现所谓“V 粒子”的明显证据，V 粒子现在被称为“奇异粒子”
1947 沙尔特岛会议召开
1948 施温格、朝永振一郎和费曼提出重正化，并计算出兰姆位移和电子的反常磁矩
1948 库什和弗利测量电子磁矩
1948 斯内尔找到中子衰变的证据，在衰变的过程中，中子会放射电子
1949 戴森给出了 QED 理论中的微扰分析
1949 费米和杨提出介子是由核子和反核子组成的
1950 比约克伦德、卡尔森和其他人发现电中性 π 介子
1950 帕诺夫斯基和其他人发现电中性 π 介子衰变成两个光子的证据
1950 宇宙射线首次提供了单次核子-核子相互作用能够产生许多强相互作用粒子的证据，过程中还会继续产生向前发射的次级粒子喷注
1950 沃德和高桥（1957）提出了沃德-高桥恒等式以及格林函数间的对称关系
1950 在研究宇宙射线的过程中首次发现强子-强子对撞产生喷注的证据
1950 萨拉姆和其他人解决了 QED 理论中重复发散的问题
1951 人们在宇宙射线中首次发现 Λ 粒子——一种奇异重子是由 uds 夸克组成的证据
1952 海森堡预言强子总散射截面会随能量增加而增加
1952 费米发现第一种共振子——Δ^{++} 粒子，它是一种由 uuu 夸克组成的重子
1952 第一台质子同步加速器——3 吉电子伏的布鲁克海文质子同步加速器 Cosmotron 开始运行
1952 欧洲粒子物理实验室（CERN）成立
1952 利文斯通、古朗特和斯奈德设计出粒子束的交互梯度调焦，并在 1959 年将其应用于 CERN 的质子同步加速器
1952 格拉泽发明气泡室
1952 派斯提出奇异粒子的关联产生
1953 博内蒂和其他人首次观察到 Σ^+ 粒子和 Σ^- 粒子，它们分别是由 uus 夸克和 dds 夸克组成的奇异粒子
1953 由 dss 夸克组成的奇异粒子——Ξ^- 粒子在宇宙射线实验中被发现
1953 盖尔曼和西岛和彦引入奇异量子数，从而预言了 Σ^0 粒子和 Ξ^0 粒子的存在

1953 斯特科伯格和彼德曼发现重正化
1954 杨振宁和米尔斯把规范性原理拓展到非阿贝尔对称性上
1954 盖尔曼和洛用重正化群研究 QED 理论
1954 伯克利质子加速器 Bevatron 投入使用，它是一台 6 吉电子伏的质子同步加速器
1955 人们在布鲁克海文质子同步加速器 Cosmotron 中发现了 Σ^0 粒子
1955 张伯伦和西格雷在 Bevatron 上发现反质子
1956 张伯伦和西格雷观测到反质子在物质中发生湮灭
1956 Bevatron 发现反中子
1956 莱因斯和寇文首次发现自由中微子和反中微子
1956 霍夫施塔德和迈利斯特发现质子不是点状粒子
1956 坂田昌一提出一种模型，认为强子由 p 粒子、n 粒子和 Λ 粒子组成
1956 内山菱友把杨-米尔斯理论推广到其他群
1956 李政道和杨振宁提出宇称在弱力中可能遭到破坏
1957 吴健雄观察到弱力衰变中宇称不守恒，这说明弱力可以分辨出左和右
1957 杜布纳 10 吉电子伏质子同步加速器投入使用
1958 古德哈伯发现电子中微子是左旋的
1959 阿尔瓦雷茨和其他人一道发现 Ξ^0 粒子，它是一种由 uss 夸克组成的奇异重子
1959 池田峰夫和其他人一道提出强子的 SU（3）对称
1959 DESY 建立
1959 雷琪提出雷琪理论，它是一种描述强相互作用散射数据的框架
1959 28 吉电子伏质子同步加速器在欧洲粒子物理研究中心 CERN 投入使用
1960 33 吉电子伏布鲁克海文 AGS 质子同步加速器投入使用
1960 南部阳一郎强调真空的对称性破缺是粒子产生的源头
1960 温伯格提出费曼图的收敛定理
1960 樱井提出第一个有关强力的杨-米尔斯理论
1961 格拉肖提出存在传递弱力的中性媒介玻色子
1961 杰弗里和史蒂芬提出坡密子概念，用于描述高能弹性散射数据
1961 阿尔瓦雷茨和他的同伴发现 ω 介子
1961 哥德斯通定理把无质量的粒子、哥德斯通玻色子同量子论系统中隐藏的连续的对称性连接到一起
1961 在质子同步加速器 Cosmotron 中发现 ρ 介子
1961 在质子加速器 Bevatron 中发现 η 介子
1961 盖尔曼和尼曼提出了强相互作用粒子的对称性方案——八正法
1961 南部阳一郎提出低质量 π 子是对称性被破缺的一种表现
1961 萨拉姆和沃德提出规范原理
1961 陶歇克设计并建造了第一台电子-正电子对撞机，这台对撞机的周长 4 米
1962 盖尔曼和尼曼提出重子的十重态，并预言存在 Ω^- 粒子
1962 斯坦福线性加速器中心（SLAC）建立
1962 莱德曼和其他人在布鲁克海文发现 μ 介子中微子

1962 “强子”这个名词被提出
1963 高能物理研究所（IHEP）在莫斯科南部的索帕克夫成立
1963 费曼认识到引入鬼粒子图的必要性
1963 索科洛夫和特诺夫解释了同步加速器辐射和磁场中运动带电粒子极化之间的联系
1964 盖尔曼和茨威格提出夸克（茨威格称其为 ace）是强子的基本组成成分
1964 在质子加速器 Bevatron 和布鲁克海文的 AGS 发现 η' 介子
1964 格林伯格、韩武荣和南部阳一郎提出夸克具有颜色属性
1964 比约肯和格拉肖提出夸克的第四种味道——粲量子数
1964 克罗宁和菲奇观测到 K 介子的 CP 对称性破缺
1964 西格斯和其他人引入西格斯机制，通过它可以推导出 W 粒子和 Z 粒子的质量
1964 萨米奥斯和他的同事在布鲁克海文发现 Ω^- 粒子
1965 范尼亚金和特仑特夫发现渐近自由的线索
1965 韩武荣和南部阳一郎用八向量场（胶子）以及颜色单态的可观测粒子讨论了 SU（3）颜色对称性
1966 斯坦福线性加速器中心线性电子加速器建成（它可以用 20 吉电子伏电子轰击固定靶）
1966 南部阳一郎进一步讨论了胶子的颜色的八重态
1966 比约肯写出描述核子自旋结构的求和公式
1967 法捷耶夫、波波夫和德威特用公式表示非阿贝尔规范理论中的鬼粒子
1967 格拉肖、温伯格和萨拉姆提出电弱力理论，统一弱力和电磁相互作用
1967 电子-核子散射的古特佛里德求和公式提出
1967 索帕克夫 76 吉电子伏同步加速器在莫斯科附近投入使用
1968 查派克发明气体导线室粒子探测器
1968 范尼查诺在雷杰理论的基础上提出双共振模型，该模型为用弦描述强子铺平了道路
1969 比约肯提出轻子-核子非弹性散射的尺度不变性假设
1969 弗里德曼、肯德尔和理查德·泰勒在 SLAC 的电子-质子散射实验中找到比约肯尺度不变性的首次证据，说明点状散射中心就是夸克，这是质子内部散射中心的自旋值为 1/2 的首次证据
1969 费曼提出核子的部分子模型
1969 卡兰和格罗斯提出了自旋为 1/2 的部分子结构函数之间的关系
1969 威尔逊提出算符乘积展开假设
1969 格罗斯和卢埃林·史密斯提出了中微子-核子散射求和公式
1969 克里普洛维奇在非阿贝尔的规范理论中区分出屏蔽效应和反屏蔽效应
1970 格拉肖、里尔普洛斯和迈阿尼研究第四种夸克——粲夸克
1970 重正化群方程的卡兰-西曼吉克形式被提出
1970 南部阳一郎提出强子模型，用可以振动、旋转的弦描述强子
1970 茨默曼证明威尔逊的算符乘积展开式假设

1971 强子总散射界面上升的首次证据在瑟普可夫被观测到

1971 62吉电子伏的ISR在CERN投入使用，它能将两束质子发生对撞，是世界上第一台强子对撞机

1971 特霍夫特和威尔特曼证明了非阿贝尔规范理论，包括那些具有自发对称性破缺的理论，都可以实现可重正化

1971 威尔逊把重正化群思想应用于临界现象

1971 Drell和Yan将部分子模型应用于强子-强子对撞过程——即Drell-Yan过程

1972 SPEAR电子-正电子对撞机在斯坦福线性加速器中心竣工

1972 400兆电子伏费米实验室（主加速环）质子同步加速器开始运行

1972 波利尼、特霍夫特和其他人一道提出纬度重正化

1972 特霍夫特完成了一篇有关渐近自由的未发表文章

1972 格利波夫和里帕托夫讨论质部分子分布函数演化和尺度不变性破坏

1972 盖尔曼和弗里奇在芝加哥的一次会议上提出可观测的粒子都是颜色单态，并提出满足杨-米尔斯方程的颜色八重态向量场

1972 范·德·梅尔提出随机冷却过程，在加速器发明之前，该过程对质子-反质子对撞机中的反质子储藏起至关重要的作用。该原理早在1981年首次应用于CERN的$Sp\bar{p}S$对撞机

1972 波奇亚特、格罗斯和其他人发现，一个轻子或带颜色夸克的家族在电弱理论中也必须保持重正化

1973 巴丁、弗里奇和盖尔曼公布他们计算出中性π子衰变为两个光子，以及电子-正电子湮灭过程的比率R。这两个结果都是存在颜色属性的重要证据

1973 弗里奇、盖尔曼和鲁特威勒总结了胶子颜色八重态的优点

1973 帕里西、卡兰和格罗斯发现尺度不变性需要存在渐近自由

1973 格罗斯、威尔茨克和波利茨发现渐近自由

1973 科尔曼和格罗斯发现，只有非阿贝尔规范理论才会出现渐近自由

1973 温伯格、格罗斯和威尔茨克提出夸克禁闭

1973 CERN的ISR观测到质子-质子散射截面随着能量的增加而增加

1973 温伯格提出标准模型，该模型囊括了强力、弱力和电磁相互作用

1973 格罗斯、威尔茨克和其他人预言深度非弹性散射过程中尺度不变性会遭到破坏

1973 CERN进行的中微子散射实验首次显示出胶子和反夸克会产生核子动量的证据

1973 小林和升川提出要解释CP破坏效应必须存在更多种类的夸克

1973 在CERN观测到弱力需要涉及一种电中性的携带力粒子（所谓的弱中性流）。这种弱中性流的发现，对人们总体上接受电弱力理论和规范理论起关键作用

1973 尼尔森、奥尔森和其他人将超导体中束缚磁场的管道作为模型，提出介子中的夸克和反夸克由弦相连

1974 杰拉德和李预言粲夸克的质量

1974 布鲁克海文国家实验室的丁肇中和SLAC的里克特发现J/ψ介子，这种粒子包含一个$c\bar{c}$夸克对

1974 “11月革命”爆发

1974 乔治和格拉肖基于 SU（5）群提出大统一理论
1974 埃利斯和贾菲分别提出质子和中子自旋结构的求和规律
1974 DESY 的 DORIS 电子-正电子对撞机投入使用
1974 VEPP-2M 电子-正电子对撞机在新西伯利亚投入使用
1974 威尔逊提出格点规范理论
1974 在费米实验室的 μ 子-离子散射实验中，首次发现深度非弹性散射中的尺度不变性破坏现象
1974 特霍夫特和波利亚科夫发现非阿贝尔规范理论存在自发对称性破缺，暗示着磁单极子的存在
1974 约翰逊和其他人提出了重子的 MIT 包模型
1975 萨米奥斯和他的同事发现第一种粲重子——由 uuc 夸克组成的 Σ_c^{++} 粒子和由 udc 夸克组成的 Λ_c^{+} 粒子
1975 佩尔和 SLAC 的同事发现 τ 重轻子，说明存在第三个轻子家族
1975 SLAC 的 SPEAR 电子-正电子对撞机首次发现喷注的证据，它支持了夸克的自旋为 1/2
1975 随机冷却原理被应用到 CERN 的 ISR 上
1975 对胶子求和的 BFKL 方程被引入
1975 施夫曼和其他人提出企鹅图（1977 年由埃利斯命名）
1975 波利亚科夫和其他人提出瞬子
1975 阿贝克斯特和波利茨提出重夸克束缚态类似于电子偶素——一种有电子和正电子组成的束缚态
1975 科内尔结合库伦项和一个“线性”项，提出夸克间存在的势
1976 SLAC 的实验人员首次发现只包含一个粲夸克的介子——D^0 介子
1976 费米实验室的工作者发现在质子-铍碰撞过程中产生的 A 重子存在天然极化
1976 SLAC 公布了首次极化电子-极化质子散射
1976 CERN 的 SPS（能够将 450 吉电子伏质子打到固定靶上）投入使用
1976 特霍夫特说明真空隧道效应是由瞬子导致的
1976 曼德斯塔姆和特霍夫特提出真空是一种双超导体，是一种磁单极子凝聚物
1976 埃利斯、盖拉德和罗斯提出三喷注事件是观测胶子的一种途径
1977 莱德曼和其他人在费米实验室发现 γ 介子，它由 $b\bar{b}$ 对组成
1977 阿尔塔莱利和帕利希推导出夸克和胶子分布演化的方程
1977 胶子凝聚物被提出
1977 威尔逊夸克和错误夸克被引入格点 QCD，以抵消重复出现的夸克
1977 斯特曼和温伯格率先研究喷注
1977 帕西和奎因利用瞬子引入一种对称性破缺方案，以保持 QCD 理论中的时间反演对称性（或者等价地说是 CP 对称性）
1978 第一台利用超导加速腔的全尺度加速器——斯坦福的 SCA（超导线性加速器）投入使用
1978 DESY 的 PETRA 电子-正电子对撞机投入使用

1978 威尔茨克和温伯格引入轴子

1978 在DESY的DORIS探测器观测到电子-正电子对撞过程中出现Υ介子

1978 施夫曼和其他人提出QCD求和公式，这是一种从算符乘积展开发展而来的工具，可以计算强子性质，它还包括了真空的性质

1978 费曼和菲尔德提出碎裂方法，以描述夸克和胶子如何产生强子喷注

1978 CERN的实验工作者们用中微子散射为QCD理论的尺度不变性破坏描述提供进一步支持

1979 康奈尔的CESR电子-正电子对撞机投入使用

1979 DESY的PETRA观测到三喷注事件，该事件中第三束喷注由胶子产生

1979 温伯格为有效量子场论奠定了基础，并重新将手性微扰理论表述为一种有效理论

1979 格鲁茨等人将蒙特卡罗法引入纯规范理论的格点分析

1980 SLAC的PEP电子-正电子对撞机开始投入使用

1980 DESY的TASSO合作小组证实胶子的自旋值为1

1980 CESR发现Υ（4S）介子

1981 CERN发现第一种包含b夸克的重子——Λ_b粒子

1981 帕利希、温格特恩等人将截断近似引入处理夸克问题的格点QCD理论

1981 尼尔森和二之宫的“行不通”理论证明，如果想在格点QCD中保持某些有用的对称性，就不可避免地会引入额外的多余夸克（“种类加倍”）

1981 在PETRA的电子-正电子对撞机中首次发现了颜色相干性的信号

1981 CESR发现B介子存在的首次证据

1981 DESY实验室人员首次发现光子结构函数的信号

1981 质子-反质子对撞机$Sp\bar{p}S$在CERN开始投入使用，$Sp\bar{p}S$中对撞质子和反质子的总能量可以达到900吉电子伏。这台机器由最初的SPS改装而成，SPS可以将450吉电子伏的质子束打到固定靶上

1982 在CERN的质子-反质子对撞机$Sp\bar{p}S$上首次观测到强子-强子对撞产生喷注的明确证据

1983 CERN的UA1和UA2合作小组发现W粒子和Z粒子

1983 SLAC的实验人员第一次验证质子自旋的求和规律

1983 CERN的实验者发现束缚态核子的结构函数同自由核子的结构函数不同（EMC效应）

1983 帕利希等人首次尝试用格点计算强子质量

1984 费米实验室的Tevatron加速器作为固定靶质子同步加速器投入使用，它是第一台基于超导磁体的加速器

1984 CERN的质子-反质子对撞产生的二喷注事件为强耦合参数在高能情况下会发生“跑动”提供了证据，并验证了胶子的无质量向量属性

1985 施莱恩和英格尔曼提出坡密子具有结构函数

1986 首次探索与夸克-胶子等离子体相关的重离子实验在CERN的SPS和布鲁克海文的AGS开始进行

1987 费米实验室的 Tevatron 以质子-反质子对撞模式投入使用，其总能量可高达 1.8 太电子伏

1987 64 吉电子伏的 TRISTAN 电子-正电子对撞机在日本国家高能物理实验室开始投入使用

1987 DESY 的 ARGUS 合作小组发现 B 介子混合效应

1987 瓦罗金、施夫曼以及伊斯格和怀斯（1989 年）引入重夸克对称性

1988 UA8 首次观测到坡密子组分发生强散射的证据

1988 CERN 的 EMC 小组提出核子自旋值要比从简单夸克模型推导出来的值低得多（自旋危机）

1989 SLAC 的线性对撞机（SLC）投入使用

1989 LEP 电子-正电子对撞机在 CERN 投入使用

1989 BEPC 电子-正电子对撞机在北京投入使用

1989 CERN 和 SLAC 公布支持只存在 3 种轻中微子的首次试验结果，这说明存在 3 个家族，总共有 6 种夸克

1989 TRISTAN 的 AMY 合作小组首次发现胶子自耦合的证据

1990 第一台探测粒子束碰撞试验的硅微带顶点探测器被安装在 SLAC 的 Mark Ⅱ 实验上

1990 乔治提出重夸克理论公式

1991 CERN 确认轻中微子只有 3 种

1991 CERN 首次发现质子中反下夸克的数量超过反上夸克的数量的证据，在夸克海中存在味道不对称

1991 人们将 SLD 探测器安装在 SLAC 的 SLC 上

1992 HERA 电子-质子对撞机在 DESY 投入使用

1992 费米实验室的 E665 小组观测到首次深度非弹性散射喷注

1992 TRISTAN 的实验人员首次发现光子中存在胶子的证据

1992 安装在强子对撞机上的第一台硅微带顶点探测器，在费米实验室的 CDF 投入使用

1992 CERN 的 NA5 小组利用产生强子的关联（π 子干涉法），首次清晰地确定了相互作用区域的大小

1993 人们第一次测量出中子自旋结构函数

1993 HERA 的实验人员第一次观测到质子包含大量具有低动量分数的夸克

1994 CERN 的实验人员首次对核子的另一种自旋结构函数 g_2 进行测量

1994 费米实验室的 CDF 和 D0 合作小组发现顶夸克存在的首次证据

1994 VEPP-4M 电子-正电子对撞机在新西伯利亚投入使用

1994 一台 4 吉电子伏连续电子束加速器在弗吉尼亚州纽波特纽斯市的杰弗逊实验室投入使用

1994 Tevatron 在质子-反质子对撞中首次发现颜色相干性

1994 温格特恩用格点计算出轻强子的质量

1994 CERN 开始进行铅离子束实验，以寻找夸克-胶子等离子体

1994 穆勒和后来的知以及拉杜施金提出普遍的部分子分布
1995 HERA 测量出坡密子的结构函数
1995 费米实验室完成了重子磁矩数据表格的精确测量
1995 顶夸克的存在被证实
1995 温格特恩用格点计算出胶球的质量和衰变率
1995 HERA 对质子中的胶子分布进行了直接测量
1995 CERN 的 SMC 完成了对核子自旋结构的首次测量，上价夸克和下价夸克之间存在差别
1996 能量达 161 吉电子伏的 LEP2 投入使用，它能够产生 W^+、W^- 粒子对
1997 人们首次发现夸克质量变化的证据
1997 CERN 发现夸克-胶子等离子体的首次证据
1997 为了完成主注入器，费米实验室主加速环停止运行
1997 第一种奇异介子 $\pi_1(1400)$ 粒子在布鲁克海文被发现
1998 CDF 发现 B_c 粒子，它是最后一种普通介子
1998 费米实验室的实验人员证实，在质子的海夸克中，反下夸克的数量比反上夸克多
1998 人们首次直接证明存在 τ 中微子
1998 CP-PACS 用格点计算出轻强子谱
1998 SLC 关闭
1998 中微子振荡的迹象首次出现，表明中微子可能具有很小的质量
1999 布鲁克海文的相对重离子对撞机（RHIC）开始运转
1999 SLAC 和 KEK 的 B 粒子工厂投入使用
1999 威尔茨克等人引入颜色-味道关联
2000 LEP 关闭
2001 人们在 B 介子中观测到 CP 破缺
2001 Tevatron Ⅱ开始运行
2002 费米实验室发现双粲重子
2003 人们发现五夸克 Θ^+ 态

附录 2 小词典

为了简洁，各词条的解释牺牲了部分精确性和完整性。各词条按拼音排序。

$\sqrt{s}$ 见不变质量。

（微扰级数的）次数　order　在振幅表达式中耦合参数的指数；指数为 1 的是一次项，指数为 2 的是二次项，等等。

ALEPH　位于瑞士 CERN 的电子-正电子对撞机 LEP 上的四台探测器之一。

AMY　位于 KEK 的电子-正电子对撞机 TRISTAN 上的四台探测器之一。

BFKL 方程　BFKL equation　一种允许包含许多低能胶子的数学方程式，例如在考虑电子同质子散射的时候会用到。

B 介子　B mesons　包含 b 夸克的介子。

C　一种与电荷有关的操作，它能够将一个粒子变为它的反粒子。C 操作能够将一个电子转变为正电子，但将一个光子进行 C 操作后仍然是一个光子。

CDF 对撞机探测器　CDF　Tevatron 质子－反质子对撞机上两台探测器之一。

CERN　欧洲粒子物理实验室，位于瑞士日内瓦附近。

CESR　位于美国康奈尔的电子-正电子对撞机速器。

CPT 对称性　CPT symmetry　由粒子-反粒子变换对称性 C、左右坐标交换对称性 P，和时间反演对称性 T，构成的复合对称性。所有场论都遵循 CPT 对称性，它保证粒子和反粒子的严格对应。

CP 对称性　CP symmetry　粒子-反粒子变换对称性 C 和右旋-左旋坐标交换对称性 P 组成的复合对称性。CP 对称性可以将左旋的中微子同右旋的反中微子联系起来。

CP 破缺　CP violation　CP 对称性的破坏，可以在 K 介子和 B 介子系统中观测到这种现象；它提供了一种区分左右的独特而绝对的方法，再此基础上或许可以解释宇宙中为什么物质远远多于反物质。

D0　Tevatron 质子-反质子对撞机上的两台探测器之一。

DELPHI　位于瑞士 CERN 的电子-正电子对撞机 LEP 上的四台探测器之一。

DESY 位于汉堡的德国加速器中心。

DGLAP 演化方程 DGLAP evolution equations 描述结构函数怎样依赖于探测波长的方程。

DORIS 位于德国 DESY 的一台旧电子-正电子对撞机。

Drell-Yan 机制 Drell-Yan mechanism 一组夸克和反夸克发生湮灭，产生一组电子-正电子对，或者产生由其他某种轻子同它的反粒子同伴组成的粒子对。

D 介子 D mesons 包含粲夸克的介子。

EMC 欧洲 μ 子合作小组，位于瑞士 CERN 的一个研究 μ 子-核子散射的研究小组。

EMC 效应 EMC effect 当核子成为一个大原子核的一部分时，自由核子内部的夸克分布情况会发生改变。

GeV 一个能量单位——吉电子伏，它的大小等于 10 亿电子伏。

H1 位于德国 DESY 的 HERA 电子-质子对撞机上的一台探测器。

HERA 位于德国 DESY 的电子-质子对撞机。

ISR 交叉存储环，位于瑞士 CERN 的一台质子-质子对撞机，在 1971 年至 1984 年期间运转。

L3 位于瑞士 CERN 的电子-正电子对撞机 LEP 上的四台探测器之一。

LEAR 位于瑞士 CERN 的低能反质子环。

LEP 位于瑞士 CERN 的电子-正电子对撞机。

LHC 位于瑞士 CERN 的大型强子对撞机。

MeV 一种能量单位（兆电子伏），等于 100 万电子伏特。粒子质量都用 MeV 来表示。例如，质子的质量是 938.3MeV。

OPAL 位于瑞士 CERN 的电子-正电子对撞机 LEP 上的四台探测器之一。

OPE 见算符乘积展开。

P 宇称（Parity）。

PEP 位于美国 SLAC 的电子-正电子对撞机。

PETRA 位于德国 DESY 的电子-正电子对撞机。

PLUTO 位于德国 DESY 的 PETRA 电子-正电子对撞机上的四台探测器中的一台。

Q^2 见迁移动量的平方。

QCD 量子色动力学，一种有关强相互作用的理论。

QCD 求和规则 QCD sum rules 近似来讲是指对强子参数的非微扰限制。

QCD 真空 QCD vacuum QCD 理论中能量最低的态。

SLAC 位于加利福尼亚斯坦福的斯坦福线性加速器中心。

SLC 位于加利福尼亚斯坦福的斯坦福线性加速器。

SMC 自旋 μ 子合作小组，位于瑞士 CERN 的一个实验小组。

SPEAR 位于美国 SLAC 的电子-正电子对撞机。

SPS 位于瑞士 CERN 的超级质子同步加速器。

$Sp\bar{p}S$ 位于瑞士 CERN 的质子-反质子对撞机。

TOPAZ 位于KEK的电子-正电子对撞机TRISTAN上的四台探测器之一。

U（1）问题 U（1）problem QCD理论中的一种明显对称性［U（1）对称性］要求像质子这样的粒子应该有同伴，但实际不存在。

W粒子 W particle 也称作W玻色子，是一种较重、带电、携带弱相互作用力的粒子。

x 核子内部的夸克或胶子携带的核子动量分数。

ZEUS 位于德国DESY的HERA电子-质子对撞机上的一台探测器。

Z粒子 Z-particle 也称作Z玻色子，是一种质量很大、携带弱相互作用力的电中性粒子。

α 电磁耦合常数。

α_s 强耦合常数。

θ真空 theta-vacuum 人们假设的一种QCD理论的真正真空，由许多拓扑性质不同的真空态混合而成。

Λ参数 Lambda（Λ）parameter 一个由实验测定的数值常数，与QCD理论作用的空间大小有关，大致等于质子的大小。该尺度参数的值大约为200兆电子伏。

μ子 muon 一种带电的基本粒子，质量为105.7兆电子伏，能够从π介子弱衰变过程中产生，通常能够衰变为电子。

π子（或π介子） pion（or π meson） 一种由上夸克、下夸克和反夸克组成的介子，是原子核中将质子和中子结合在一起的强相互作用力媒介粒子。

A

阿贝尔群 Abelian group 一种关于对称性的数学表达，其中两次对称性操作，无论顺序如何，总能够给出同样的结果。U（1）群就是阿贝尔群。

暗物质 dark matter 一种看不见的物质，人们认为它占据了宇宙总质量的很大一部分。

B

靶恩 barn 一种面积单位，等于10^{-28}平方米，常用它来描述散射实验。

包含过程 inclusive process 在所有产物中只有一个粒子需要证实和测量的一种过程。余下的事件产物都被忽略掉了。

比约肯求和规则 Bjorken sum rule 质子和中子径向自旋组分的差别同弱相互作用参数之间的关系。

标记 tagging 利用碰撞碎片的一些性质作为某个特殊过程的信号。

标准模型 Standard Model 有关弱力的电弱理论同有关强力的QCD理论的统称，可以描述人们熟悉的各种基本粒子。

玻色子 boson 一种自旋为零或整数的粒子，例如光子，许多这样的粒子可以在同一时刻居于同一个态上。

不变量 invariant 在变换下不发生改变的量。

不变质量　invariant mass　在某次碰撞中能用作产生粒子的有用能量，表示为$\sqrt{s}$。不变质量对于所有观测者都具有相同的值，它是一种洛仑兹不变量。

不相容原理　exclusion principle　该原理说明，两个电子或者两个任何其他种类的费米子，都不可能占据完全相同的量子态。通常也称为泡利不相容原理。

部分子　parton　在核子中存在的散射中心，经证实是夸克、反夸克或胶子。

部分子模型　parton model　一种核子模型，人们认为核子由称为部分子的简单散射中心组成，这些部分子经证实就是夸克和胶子。同简单部分子模型不同，经过"QCD改进"的部分子模型考虑部分子之间的相互作用，它们因此不再是具有简单父核子动量分数的严格点状散射体。

C

粲量子数　charm　粲夸克或c夸克有关的一个量子数，粲夸克具有的粲量子数值为+1。

粲素态　charmonium states　由粲夸克-反粲夸克对激发态组成的一系列介子。

场　field　分布在某个区域上的量或者作用，而不是集中在某个点上的东西。

场论　field theory　任何所研究的量用场而不是用点表示的理论。在量子场论中，场是许多具有不同能量的包的集合。

超对称性　supersymmetry　将具有整数自旋的粒子同具有半整数自旋的粒子联系起来的对称性。尽管实验物理学家们至今尚未发现超对称性的证据，研究该问题的理论已经在理论物理学家当中引起了极大的轰动。

尺度不变性　scale invariance　用来测量系统的尺子长度发生变化时，测量结果不发生任何改变的性质。

尺度参数　scale parameter　见Λ参数。

重正化　renormalization　一种从量子场论中获取物理内容的方法，如果不这样做，就容易得到无穷大的无意义结果。

重正化群　renormalization group　一套用来修正理论参数，并能保持有意义的物理特征不发生改变的数学变换。

触发器　trigger　在粒子探测器中，用来判断忽略还是记录事件的决策系统。

传播子　propagator　粒子从一个点传播到另一个点的几率振幅。

磁矩　magnetic moment　在施加外部磁场之后，磁铁偏转特性的度量。

磁振子　magnon　同磁性物质中磁化相关的无质量的量子粒子。

存储环　storage ring　一种粒子加速器，该机器中的粒子束会绕大致相同的圆形路径转很多次，特别是在增加粒子数量或者存储以后要用的粒子束时很有用。

D

大爆炸　Big Bang　人们假设宇宙诞生于一个巨大的火球，随后经历了上百亿年的冷却和膨胀。

单极子　monopole　顾名思义，单极子就是一个单独的极。磁单极子是一个单独的磁北极或磁南极，但这些磁单极都没有在大自然中出现过。QCD理论中的单极子是一种色-磁场单极，例如在夸克禁闭的双超导模型中出现的情形。

单态　singlet　只有一个成员的“家族”，该成员中，人们感兴趣的组分特征以一种特殊方式结合，形成一个值为零的单态。例如，一个中性原子是电荷单态，像Λ粒子这样的同位旋单态同位旋为零。可观测到的参与强相互作用的粒子都是颜色单态，这说明它们虽然都由夸克组成，但它们总体上不带剩余颜色荷。

弹性散射　elastic scattering　粒子互相“弹开”的散射过程，该过程产生的粒子和碰撞前相同。

氘　deuterium　氢的一种重同位素，它的原子核称氘核，由1个质子和1个中子组成。普通的氢核只有1个质子。

底量子数　bottom　一个与底夸克或称b夸克有关的量子数，b夸克所带的底量子数为-1。它有时也被称为美（beauty）量子数。

底素态　bottomonium states　由底夸克-反底夸克对的激发态组成的一系列介子。

电弱理论　electroweak theory　有关弱力的格拉肖-温伯格-萨拉姆理论，根据该理论，弱相互作用通过W玻色子和Z玻色子传递。

电子　electron　一种稳定的点状基本粒子，是基本电荷单位的携带者。因为电子自旋为+1/2，所以它是费米子，其质量为0.511兆电子伏。原子核周围都环绕着轨道电子。

电子伏　electronvolt　一种非常小的能量单位。

电子偶素　positronium　由电子和正电子组成的一种寿命极短的“原子”。

动夸克　dynamical quarks　在格点QCD中，那些从真空中产生的夸克。它们与计算一开始就存在的价夸克不同。

多态　multiplet　满足某种对称性的一些粒子的总称。

F

发光度　luminosity　加速器中粒子束的亮度。

发散　divergence　在计算中出现无穷大的结果，可能是由于一些无限制增长的量或者除以一个趋于零的量引起。

法捷耶夫-波波夫鬼粒子　Faddeev-Popov ghost　见鬼粒子。

反常磁矩　anomalous moment　粒子的磁矩值与理想点状粒子磁矩的偏离。

反粒子　antiparticle　与正常粒子具质量相同电荷相反的粒子。正电子是电子的反粒子伙伴；像光子或π^0粒子等这样的粒子是它们本身的反粒子。

非阿贝尔群　non-Abelian group　一种对称性的数学表达，在这种对称性中，两个对称性操作的顺序不同会导致不同的结果。SU（2）和SU（3）群都是非阿贝尔群。

非弹性散射　inelastic scattering　一种散射过程，该过程中碰撞粒子的能量被用来产生新粒子，因此离开碰撞点的粒子同入射粒子并不完全相符。

非禁闭转变　deconfinement transition　从一个夸克-胶子禁闭的态到一个不禁闭

的态的跳变过程。人们假设在高温和高密度条件下会发生这种跳变。

非微扰 non-perturbative 表示无法用一种称为微扰论的近似技巧分析的一些性质或过程。QCD理论的非微扰区域是低能、长程区域，例如在夸克结合形成质子时遇到的情形。

费曼图 Feynman diagram 在对某些粒子的物理过程进行近似数学计算时，某一部分的图像表示。

费米子 fermion 一种具有半整数自旋的粒子，像电子、质子或夸克等。费米子遵循不相容原理。

分解光子 resolved photon 一种通过与之相关联的夸克和胶子发生相互作用的光子。粗略地讲，是一种碎裂的光子。

辐射修正 radiative corrections 在计算某些过程发生几率的过程中，发射某些额外粒子会引起辐射修正，例如电磁过程中会发射低能光子。

G

高纠缠效应 higher-twist effects 由粒子之间的相关性引起的对粒子相互作用的贡献，这种效应在低能区域内变得更加显著。这些都是无法准确计算的非微扰效应。

哥德斯通玻色子 Goldstone bosons 一种产生于自发对称性破缺的无质量的粒子。

格林函数 Green's function 在场论中的某个物理过程的振幅，它是一些列费曼图的叠加。

共振子 resonance 一种不稳定的粒子，能够通过强相互作用很快衰变为其他发生强相互作用的粒子。

孤波 soliton 一种不会扩散和衰退的孤立波动。

固定靶实验 fixed-target experiments 用入射粒子束轰击静止靶——例如一块铁的实验，最终产生的碎片会分布在入射粒子方向的一个圆锥角内。

光电倍增管 photomultiplier tube 能够将一种微弱闪光变成电脉冲的探测设备。

光致产生 photoproduction 同真实或近似真实的光子发生的散射。

光子 photon 一种光“粒子”；换言之，是一种电磁能量子。

广义相对论 general relativity 爱因斯坦的引力理论，引力和时空的几何性质相关。

规范场 gauge field 为了保持局域对称性而引入的一种场。在量子理论中，它们都是自旋为1的传递力的玻色子；光子、W粒子和Z粒子以及胶子都是规范场。

规范对称性 gauge symmetry 在场的相位中存在的与时空有关的对称性。

规范化 regularization 将无穷消除掉的数学过程。

规范理论 gauge theory 一种相互作用由局域对称性质决定的量子理论。

硅顶点探测器 silicon vertex detector 大多数粒子探测器的内层，它依靠硅技术定位距粒子原始相互作用点非常近的粒子轨迹分叉。

硅微带探测器 silicon microstripe detector 硅顶点探测器的基本元件，它依靠安放在另一块半导体中的半导体带探测路过的粒子。

鬼粒子　ghost　一种对特定费曼图的贡献；鬼粒子对应着非阿贝尔规范理论中保持几率守恒需要的非物理粒子。

H

核子　nucleon　中子和质子的统称。

红外发散　infrared divergence　一种与能量趋于零有关的发散。

J

基态　ground state　量子力学系统中能量最低的态。

极化　polarization　粒子自旋相对于它前进方向的指向。例如在右手系中，右极化的粒子其自旋方向同前进方向一致。在横向极化中，粒子自旋方向同它飞行的路径成直角。

价夸克　valence quarks　那些携带父粒子量子数的夸克。

准哥德斯通玻色子　pseudo-Goldstone bosons　低质量粒子，如 π 子，起源于并不严格的对称性自发破缺。

简单部分子模型　naive parton model　部分子模型的简单版本，该模型中每个点状部分子携带父核子动量的一部分，并满足比约肯尺度不变性。

渐近自由　asymptotic freedom　夸克之间的颜色力随着它们之间的距离的减小而减小。

交错夸克　staggered quarks　在格点 QCD 中，可以通过将夸克信息分布在几个格点位来消除偶然产生的多余夸克，该技术方法产生的夸克称为交错夸克。

交替梯度同步加速器　AGS　美国布鲁克海文国家实验室的一台质子加速器。

胶球　glueball　一种完全由胶子组成的中性介子。

胶子　gluon　根据 QCD 理论，胶子携带强相互作用力（强力）。胶子是一种基本的、无质量、电中性、自旋为 1 的玻色子。它带有 8 种不同的颜色。在费曼图中，它一般用螺旋状曲线表示。

结构函数　structure functions　核子组分的动量分布。

介子　meson　一种发生强相互作用的粒子，由夸克和反夸克组成，例如 π 介子。

禁闭　confinement　QCD 理论的一项关键假设：自由夸克和胶子不能够存在，而必须居住在像质子或中子这样的复合粒子当中。这些粒子中的总颜色最终混合为白色，用技术术语来说就是颜色单态。

纠缠　twist　见高纠缠效应。

局域变换　local transformation　在时空内的每一点都各不相同的变换。

K

卡西米效应　Casimir effect　由于真空扰动引起的、能将两块导体板向一起推的力，这是零点能的一个结果。

库珀对　Cooper pairs　在超导体中传导电流的物体；由两个松散的电子对组成的非局域自旋为1的粒子。

夸克　quark　参与强相互作用的基本粒子，颜色荷的携带者。此外，夸克还带有电荷，并参与电弱相互作用。存在6种类型（或者称味道）的夸克：上（u）夸克、下（d）夸克、奇异（s）夸克、粲（c）夸克、底（b）夸克和顶（t）夸克。

夸克-胶子等离子体　quark-gluon plasma　一种由“自由”夸克和胶子组成的气体。

夸克凝聚态　quark condensate　一种从真空中出现的由夸克-反夸克对组成的海，它形成了强相互作用的背景。

夸克素　quarkonium　一种由重夸克和反夸克对组成的介子。复数形式为 quarkonia。

宽度　width　粒子趋向衰变的量度；它是粒子寿命的倒数。宽度窄意味着粒子寿命长。

L

雷杰理论　Regge theory　用交换介子家族表示强相互作用散射的理论。

李群　Lie group　一种对应着连续、平滑变换的对称群，例如转动变换。

粒子束管道　beam pipe　在加速器中，位于最里面的包含粒子束的管道。

粒子束碰撞试验　colliding beam experiments　一束粒子同另一束粒子迎面相撞的实验。相对于固定靶实验，其碰撞发生的可能性很小，但可以得到特别高的碰撞能量。

粒子数据小组　Particle Data Group　一个由粒子物理研究者组成的国际小组，成员们共同维护和更新 *Review of Particle Properties*（《粒子性质评论》）——这是粒子性质以及其他关键粒子物理信息的最终列表（见 http://pdg.lbl.gov）。

量子场论　quantum field theory　量子力学的最高形态。在该理论中，物质和力都用量子波求和得到的场来描述。量子场论可以自然地允许粒子产生和湮灭。

量子数　quantum number　用来标记或认证量子态的简单数字或向量以及在该态进行测量得到的结果。守恒定律都用量子数来表达。

临界点　critical point　表明两相系统在哪里出现、在哪里变得无法区分的条件，例如温度和压强。

临界现象　critical phenomena　两相物质之间的区别消失的现象。

零点能　zero-point energy　残余的最小能量，在此能量之下的量子态系统无法存在。

流夸克质量　current quark mass　与无质量强相互作用理论所具有的对称性发生破缺有关的小夸克质量。

路径积分　path integral　量子态之间跃迁的振幅，表现为联系这些态之间所有可能路径的总和。

M

迈斯纳效应　Meissner effect　在超导体中，超导材料内部产生的磁场排斥效应。

麦克斯韦方程组　Maxwell's equation　由詹姆士·克拉克·麦克斯韦（James Clerk Maxwell）提出的有关电和磁的方程组。

N

凝聚态　condensate　一种态，当许多相同的粒子通过某种形式互相关联，形成的一种总能量最低的结构。夸克、单极子和库珀对都能够形成凝聚态。

O

偶产生　pair production　在QED理论中，一个高能光子湮灭，产生一组电子-正电子对的过程。

耦合"常数"　coupling "constant"　一个用来表达相互作用强度的无量纲数字。微扰论结果可以用耦合常数的次幂表示。通常，"参数"这个术语比"常数"更加恰当。

P

排除过程　exclusive process　一种散射过程，在该过程中，所有末态粒子都需要探测和测量。

喷注　jet　一种具有一定方向、包含强相互作用粒子的圆锥形粒子束。

皮靶恩　picobarn　散射截面的一种单位。1皮靶恩等于10^{-12}靶恩，或者等于10^{-40}平方米。

漂移室　drift chamber　用来描绘粒子轨迹的粒子探测器元件；主要依靠带电粒子经过时使气体发生电离进行测量。

平均寿命　mean life　一个孤立粒子在湮灭前能够存在的平均时间。它是衰变率的倒数，与半衰期——与粒子衰变几率为50%相对应的时间——成正比。半衰期的值大约为平均寿命的70%。

坡密子　Pomeron　一种（假定的）传递某种高能强相互作用的交换粒子。主要由胶子组成。

普遍性　universality　不考虑颜色和味道属性，夸克和胶子之间的基本耦合强度完全相等。

Q

奇异数　strangeness　奇异粒子所具有的一种性质，起源于奇异夸克或反奇异夸克的存在。奇异夸克所具有的奇异数是-1。

气泡室　bubble chamber　一种粒子探测器。气泡室中，粒子的轨迹通过过热液体中的气泡径迹表现。

迁移动量的平方　momentum transfer squared　传递相互作用的虚粒子的不变质量（的相反数），用 Q^2 表示。例如一个电子质子散射实验，Q^2 表示传递散射的虚光子。这个量的值大就表示探测波长短，就能够探测到更细微的内容。

强 CP 问题　strong CP problem　在 QCD 理论中，为什么瞬子产生项的参数值很小。

强子　hadron　发生强相互作用的粒子。包含 3 个夸克的强子称为重子，包含夸克-反夸克对的强子称为介子。

轻子　lepton　像电子和 μ 子这样的物质粒子，它们不参与强相互作用。

轻子数　lepton number　一个在所有反应中都守恒的量，等于轻子的数量减去反轻子的数量。电子、μ 子和 τ 子的轻子数各自守恒。

求和规律　sum rule　核子中夸克或胶子总动量分布之间的关系，或由此而导致的值。

全局变换　global transformation　一种对每个时空点都相同的变换。

群　group　对称性的一种数学表达。在形式上，数学中的群是一些遵守 4 种特殊规律的元素的集合。

R

热量计　calorimeter　一台粒子探测器的外侧部分，它能够探测由相互作用产生的粒子的能量。

韧致辐射　bremsstrahlung　带电粒子发射光子而损失能量，或者是夸克（或反夸克）发射胶子而损失能量。

软相互作用　soft interaction　一种迁移动量很小的相互作用。

S

散射　scattering　一束粒子或辐射同某个靶之间的相互作用。粒子束一般会改变方向，或至少有一部分粒子束改变方向。

散射截面　cross-sectiong　一种控制两粒子相互作用可能性的特殊面积值。

色磁力　chromo-magnetic force　两个运动色荷（特别是具有自旋的颜色荷）之间的颜色力，是运动电荷之间磁力的颜色类比。

色电力　chromo-electric force　两个静止色荷之间的颜色力，是两个静止电荷之间静电力的颜色类比。

闪光物质　scintillator　一种用于粒子探测器的材料，在粒子经过时，它会产生相应的微弱闪光。

深度非弹性散射　deep inelastic scattering　用来刺探靶核子最内部组分的电子（或者 μ 子、中微子）发生的散射过程。

声子　phonon　固体中存在的一种振动能量包（量子）。

势能　potential　一般指包含力的系统变化时导致的可用能量。势能随距离的变化就是力场的相对强度。在量子场论中，势变为了携带力的量子。

手性　handed　左旋和右旋之间的一种差别。普通的木螺钉都是右旋的，因为它在顺时针转的时候会在木头中向前运动。

手性对称性　chiral symmetry　与左旋和右旋无关的对称性。

手性恢复　chiral restoration　人们推断这种手性对称性的恢复会在高温高密条件下发生。

手性微扰论　chiral perturbation theory　一种相对于普通 QCD 计算而言的非微扰技术，其中，QCD 理论的手性对称性破缺部分被看做是满足手性对称性的小微扰。

守恒定律　conservation laws　描述在某些过程中保持不变的量的规律。最常见的例子就是能量守恒和动量守恒。在粒子物理当中，也存在许多守恒定律，例如强相互作用中的重子数守恒和同位旋守恒。守恒定律联系着用来描述和区分相互作用的守恒量以及相互作用的基本对称性。

衰变率　decay rate　一个独立粒子在衰变前平均寿命的倒数。

瞬子　instanton　一种起源于拓扑性质的、短暂的无质量胶子“结”，它能够改变夸克的手性。

斯拉夫诺夫-泰勒恒等式　Slavnov-Taylor identities　在非阿贝尔规范理论中振幅之间的数学关系，可概括为沃德恒等式。

四矢量　four-vector　用一组四个数字来表示一个量，例如时空间隔。四矢量的大小在洛伦兹变换下保持不变。

四维动量　four-momentum　在普通的力学动量（“三维动量”）的基础上进行扩展，包含额外的第四部分——能量。

算符乘积展开　operator product expansion　用非微扰背景特征表示某些非微扰过程的短程行为，这些背景特征的相对贡献可以用微扰论计算。

随机冷却　stochastic cooling　对带电粒子束（特别是反质子）进行挤压，产生一束适合进一步加速的致密粒子束。

碎裂　fragmentation　夸克或胶子变成可观测到的、发生强相互作用的粒子的过程。

碎裂函数　fragmentation function　一个夸克或胶子碎裂成一个介子的可能性，该介子的动量分数等于介子的动量除以原始夸克或者胶子的动量。

T

通道　channel　某些过程发生的几条路径之一。

同步加速器　synchrotron　一种加速脉冲的频率同磁场的强度同步增长，以保持加速粒子运动在圆形路径上的加速器。

拓扑　topological　一种与几何性质无关的总体特征。橄榄球和足球在拓扑上都是一样的（都是球），但在几何上是不同的（只有一个是规则球形）。

W

万亿电子伏加速器 **Tevatron** 美国费米实验室的质子-反质子对撞机。

威尔逊夸克 **Wilson quarks** 在格点规范理论中手性对称性破缺的夸克，它的多余项会被格点主方程中附加的“威尔逊项”吸收。

微扰理论 **perturbation theory** 一种数学技巧，用它可以将很难处理的问题表示为一系列重要性递减的项，从而得到近似答案。

维度规范化 **dimensional regularization** 利用连续变化的维度参数将费曼图闭圈引起的无穷大孤立出来。

味道 **flavor** 夸克的类型属性。6种味道分别是上、下、奇异、粲、底和顶。

温伯格-萨拉姆理论 **Weinberg-Salam theory** 电弱理论的别称。

沃德-高桥恒等式 **Ward-Takahashi identities** 推广的沃德恒等式。

沃德恒等式 **Ward identity** 振幅之间的一种关系，将量子场论的重正化时非常有用。有时也称为沃德-高桥恒等式，在非阿贝尔情形称为斯拉夫诺夫-泰勒恒等式。

物质波 **matter wave** 同电子这样的物质“粒子”相关联的量子波动。

X

西格斯粒子 **Higgs particle** 一种已经提出但尚未观测到的粒子。在关于弱力的电弱理论中，在对称性破缺的出现以及粒子质量的起源方面具有重大作用。

狭义相对论 **special relativity** 可以应用到运动速度接近光速的物体上的力学。

线性加速器 **linac** 一种直线加速器，其电极都排列在一条直线上（因为是线性的），这样被加速的粒子在效果上就像在机器那么长的电磁波上冲浪。

相 **phase** ①波动过程相对于某些起点的度量。②某些物质的物理态，例如水的液相和气相，或者铁的磁化相和非磁化相。

相变 **phase transition** 系统从一个相到另一个相的变化。例如水烧开后便成为水蒸气，或者水蒸气凝结成水。

相干 **coherence** 两列波之间的一种关系。两列具有相同频率的波，如果一列波总比另一列波落后相同的量，就称这两列波相干。

相位稳定性 **phase stability** 相对论性的加速粒子之间的相对时间以及加速电压保持不变。

协变 **covariance** 物理量在进行某些变换（例如洛伦兹变换）之后，它们之间的关系保持不变的性质；也称为形式不变性。

信号 **signature** 从粒子散射实验中得到的，能够标志发生某种人们感兴趣的特殊过程的结果。

Y

颜色 **color** 夸克和胶子所具有的一种性质，这种性质是它们之间强相互作用的

源头，特别类似电荷是带电粒子之间点相互作用的源头。存在红、绿和蓝 3 种不同的颜色。

颜色单态　color singlet　一种带颜色的物体的组合，像 3 个夸克组成的质子，其所带的总颜色荷为零。可观测到的粒子都是颜色单态。

颜色相干　color coherence　连续发射的低能胶子之间破坏性的干涉现象，它会导致发射胶子具有角度序列以及三喷注事件中夸克喷注和反夸克喷注间粒子的抑制。

衍射散射　diffractive scattering　一种散射过程，其中的一个粒子发生弹性散射，保持不变，而另一个粒子发生破裂。

幺正变换　unitary　一个用来描述粒子家族变换的术语，这种变换保持总体粒子成分不变。与钟表的指针转动时，指针长度保持不变类似。

幺正性　unitarity　碰撞过程中总几率守恒。

硬相互作用　hard interaction　一种动量传递很大的相互作用。

有效量子场论　effective quantum field theory　应用在特定能量范围上的一种非重正化量子场论。

宇称　parity　镜像变换的数学形式。在宇称操作下，一个左旋粒子会变成右旋粒子，反之亦然。

云室　cloud chamber　一种老式粒子探测器，其中可观测到粒子的轨迹。粒子经过蒸汽时会产生由小液滴组成的“蒸汽轨迹”，从而可以追踪粒子的行踪。

Z

真空　vacuum　在日常意义上，就是指将密封盒子中的空气抽空后留下的虚无。在量子场论中，是指能量最低的态。

真空极化　vacuum polarization　像光子这样的粒子，通过扰动周围的真空而产生的能量。

正电子　positron　电子的反粒子同伴，具有一个单位的正电荷。

直接光子　direct photon　一种光子，它作为整个光子与靶的带电组分之间相互作用，而不不需要先分解。

质子　proton　一种参与强相互作用、带正电的粒子，夸克组分为 uud。氢原子核就是质子。原子核都是由质子和中子组成。

中微子　neutrino　一种质量很小、发生弱相互作用的物质粒子，不参与强相互作用。有 3 种类型的中微子存在。

中子　neutron　一种发生强相互作用的中性粒子，夸克组成为 udd。原子核就是由中子和质子构成的。

重子　baryon　一种质量很大、发生强相互作用、由 3 个夸克组成的物质粒子，例如质子。

重子数　baryon number　一个在所有反应中都守恒的量，其值等于重子的数量减去反重子的数量。

轴子　axion　一种相互作用很弱的粒子，现在假定它具有很小的质量，引入它是

为了解决QCD理论中的一个对称性问题。人们从未观测到它，它也是宇宙中丢失质量的一个候选者。

主方程　master equation　QCD理论的主方程是一套与夸克、胶子以及它们之间基于颜色的相互作用有关的能量方程。将其输入到费曼的路径积分当中能够产生费曼图。对于那些了解这方面内容的人来讲，主方程最终是一种拉格朗日密度。

自发对称性破缺　spontaneous symmetry breaking　当能量最低态的对称性少于相互作用的对称性时，对称性减少的现象。

自旋　spin　基本粒子的一种属性，同日常物体的自旋相类似，例如自旋向上。基本粒子的自旋只能取某些特定的值。具有整数自旋（单位是$\hbar$）的粒子叫做玻色子，具有半整数自旋的粒子称为费米子。

总发光度　integrated luminosity　加速器中粒子束的亮度对事件的积累。

族　families　基本“物质”粒子，或称费米子，可以分为3套对应的粒子，其中第一套包括上夸克、下夸克、电子和电子中微子。

组分夸克质量　constituent quark mass　通过夸克模型得到的夸克质量；上夸克和下夸克的组分质量大致都相当于质子质量的1/3。

最小作用量原理　principle of least action　规定力学系统必须以作用量取最小可能值的方式演化的规律。

作用量　Action　动能和势能之差对时间的求和。

第一推动 FIRST MOVER

你能感觉和触摸到的世界是由原子构成的，原子是能够分辨的最小物质块。但是原子中心本身又是一个全新的世界，其中的居民是夸克：夸克看不到，不可思议地小，但却是构成我们这个宇宙的最小砖块。在夸克统治的这个世界中，规则与我们的世界大不相同。这些规则是量子规则。巨大的粒子加速器可以将这个世界展现在物理学家们眼前，使他们能够形成一套关于量子规则的理论，用来解释夸克如何感觉彼此的存在。《量子夸克》讲的就是这套理论：量子色动力学。

上架建议：科普类

ISBN 978-7-5357-5236-9

9 787535 752369 >

定价：39.5 元